

Chapter 2

Consonants

2.1 Introduction

Bangla consonants are broadly classified into plosives, affricates, fricatives, lateral, taps, and trill according to their manners of production. Again in each of these manner-based categories there could be categorization on the basis of the place of articulation. All these traditional categorizations of consonants in SCB are shown in Table 2.1 with example words for consonants in initial, medial and final positions, as are available in literatures (Chatterji 1926; Hai 1989).

The aforesaid categorization was done by expert phoneticians and linguists.

However to suit the need of technology, objective quantifiable acoustic parameters as exists in today's speech sound is imperative. Fortunately, in recent times, with the advent of technology it is possible to objectively and quantifiably determine the manners and places of articulation of a particular consonant uttered a number of times by a number of speakers. Statistics allows us methodology to select the appropriate representative. Such objective categorization is necessary for the development of speech technology like ASR and TTS engines. As has been said earlier in India this type of activity began in early 70s (Datta et al. 1974) in Kolkata with the procurement of the first Sonagraph from Kay Elemetrics. However the work on Bangla started much later.

2.2 Experimental Procedure

Some of the instrumental approaches like signal processing tools including extraction of spectral structures have been in use in India since 70s. The use of spirometers, electroglottograms (EGG), palatographs, etc., is also quite old. A new entrant in this field is the electropalatograph (EPG). This is now in use in India. While the spirometers and electroglottograms are used for detecting the manner

Table 2.1 Place and manner of articulation of the Bangla consonant

S. No.	Grapheme	IPA symbol	Example Bangla words with phoneme in			Place of articulation	Manner of articulation
			Initial	Medial	Final		
1	ক	/k/	/kobite/- 'Poetry'	/kake/ 'Uncle'	/obek/ 'Surprise'	Velar	Unaspirated unvoiced stop
2	খ	/kʰ/	/khotom/ 'End'	/akhil/ 'Whole'	/ekʰ/ 'Sugar cane'	Velar	Aspirated unvoiced stop
3	গ	/g/	/gedhe/ 'Donkey'	/agedʰ/ 'Plenty'	/ʃʱag/ 'Goat'	Velar	Unaspirated voiced stop
4	ঘ	/gʱ/	/ghor/ 'Home'	/agʱat/ 'Injury'	/bagʱ/ 'Tiger'	Velar	Aspirated voiced stop
5	ঙ	/ŋ/		/kəŋkal/ 'Skeleton'	/baŋ/ 'Frog'	Velar	Nasal murmur
6	চ	/tʃ/	/tʃel/ 'Rice'	/əʃʃil/ 'Still'	/ʃʱemof/ 'Spoon'	Post-alveolar	Unaspirated unvoiced affricate
7	ছ	/tʃʰ/	/tʃʱegol/ 'Goat'	/biʃʱene/ 'Bed'	/meʃʱ/ 'Fish'	Post-alveolar	Aspirated unvoiced affricate
8	জ	/dʒ/	/dʒol/ 'Water'	/adʒkel/ 'Now a days'	/kadʒ/ 'Work'	Post-alveolar	Unaspirated voiced affricate
9	ঝ	/dʒʱ/	/dʒʱot/ 'Storm'	/medʒʱel/ 'in middle of'	/sādʒʱ/ 'Evening'	Post-alveolar	Aspirated voiced affricate
10	ঞ	/ɲ/		/goʃei/ 'vaisnava guru'		Post-alveolar	Nasal murmur
11	ট	/ʈ/	/ʈok/ 'Sour'	/ʈiʃke/ 'Fresh'	/mot/ 'Total'	Palatal	Unaspirated unvoiced stop (retroflex)
12	ঠ	/ʈʰ/	/ʈʱekur/ 'idol'	/kātʱel/ 'Jack fruit'	/kəʈʱ/ 'Wood'	Palatal	Aspirated unvoiced stop (retroflex)
13	ড	/d̪/		/ed̪e/		Palatal	Unaspirated voiced stop (retroflex)

(continued)

Table 2.1 (continued)

S. No.	Grapheme	IPA symbol	Example Bangla words with phoneme in			Place of articulation	Manner of articulation
			Initial	Medial	Final		
			‘Pulses’	‘Gossip’			
14	ঢ়	/dʱ/	/dʱel/ ‘Shield’	/ɔdʱel/ ‘Plenty’		Palatal	Aspirated voiced stop (retroflex)
15	ঞ	/ɲ/		/ʲeɲdʱe/ ‘Cold’		Palatal	Nasal murmur
16	ত	/t/	/tumi/ ‘You’	/betesʱ/ ‘Wind’	/otit/ ‘Past’	Alveolar	Unaspirated unvoiced stop
17	থ	/tʰ/	/tʰekʱ/ ‘Stay’	/kʱatʰe/ ‘Talk’	/pʱatʰ/ ‘Road’	Alveolar	Aspirated unvoiced stop
18	দ	/d/	/doi/ ‘Curd’	/kodel/ ‘Spade’	/pʱad/ ‘leg’	Alveolar	Unaspirated voiced stop
19	ধ	/dʱ/	/dʱhen/ ‘Paddy’	/edʱer/ ‘Container’	/bʱadʱ/ ‘Kill’	Alveolar	Aspirated voiced Stop
20	ন	/n/	/ne/ ‘No’	/onek/ ‘Many’	/bon/ ‘Woods’	Alveolar	Nasal murmur
21	প	/p/	/pʱekʱ/ ‘Ripe’	/kʱapʱ/ ‘Forehead’	/pʱap/ ‘Sin’	Bilabial	Unaspirated unvoiced stop
22	ফ	/fʰ/	/phol/ ‘Fruit’	/epʰfoʃ/ ‘Repentance’	/sepʰ/ ‘Clear’	Bilabial	Aspirated unvoiced Stop
23	ব	/b/	/boi/ ‘Book’	/kobite/ ‘Poetry’	/ʃʱb/ ‘corpse’	Bilabial	Unaspirated voiced stop
24	ভ	/bʱ/	/bʱoy/ ‘Fear’	/ebʱesʱ/ ‘Hint’	/ʳebʱ/ ‘Profit’	Bilabial	Aspirated voiced stop
25	ম	/m/	/me/ ‘Mother’	/aner/ ‘My’	/ʱam/ ‘Mango’	Bilabial	Nasal murmur

(continued)

Table 2.1 (continued)

S. No.	Grapheme	IPA symbol	Example Bangla words with phoneme in			Place of articulation	Manner of articulation
			Initial	Medial	Final		
26	ব	/t/	/rokʈo/ 'Blood'	/kore/ 'Do'	/tomor/ 'Your'	Alveolar	Trill
27	ক	/l/	/lel/ 'Red'	/belok/ 'Boy'	/kel/ 'Time'	Alveolar	Lateral
28	খ, ঞ	/ʃ/	/ʃole/ 'Stick'	/eʃi/ 'Eighty'	/ʃoleʃ/ 'A flower'	Post alveolar	Fricative
29	স	/s/	/set/ 'Seven'	/baestʃo/ 'Busy'	/tes/ 'Playing cards'	Alveolar	Fricative
30	হ	/h/	/het/ 'Hand'	/bihar/ 'To travel'	/bah/ 'Wah'	Glottal	Fricative
31	ঢ়	/tʃ/		/boʈo/ 'Big'	/ʃot/ 'Conspiracy'	Palatal	Unaspirated flap (Retroflex)
32	ড়	/tʃʰ/		/driʈʰo/ 'Rigid'	/bœʈʰ/ 'Flood'	Palatal	Aspirated flap (Retroflex)
33	ষ	/j/		/pejɔ/ 'Leg of a table'	/hoj/ 'Is'	Palatal	Approximant
34		/w/			/hewæ/ 'Wind'	Bilabial	Approximant

based categorization the palatographs are used for determining the place of articulation, where tongue is the moving articulator. Video capture is used primarily for determining lip closure. The advancement in spectral processing and interpretation has made use of spirometers and EGG important for pathological investigation, though has become replaceable for acoustic phonetic analysis for two obvious reasons. One is that wearing the electrodes is not only inconvenient but also makes the normal informants somewhat tense. Therefore one cannot really examine acoustics of free normal speech. The other is that spectral processing is very reliable, fast and can be done offline even on previously recorded speech.

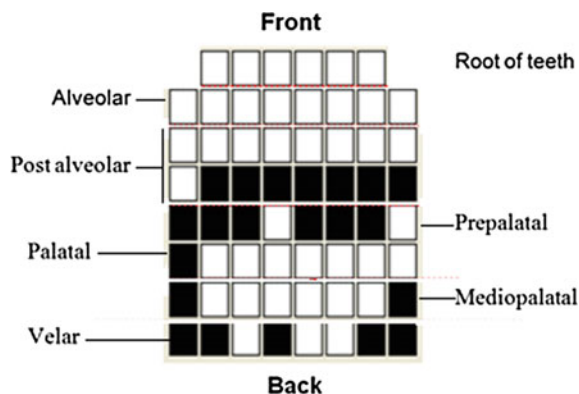
In the present report Electropalatograph (EPG) in conjunction with video camera has been used to determine the place of articulation of a large number of consonants. For acoustic signatures spectral structure has been used. The EPG system is used to objectively ascertain the place of contact made by the tongue with the hard palate and video camera is used to determine if the closure is made at the lips. For extracting the spectral information commercially available soft-ware packages like Cool-Edit Pro and Wave surfer has been used.

The EPG consists of a custom made palate to fit the informant. This artificial palate extends from the root of the teeth only up to the beginning of the soft palate. The extension of the palate deep into the palate to cover velar region fully is avoided as it would otherwise cause discomfort to the wearer even may do harm to the soft palate. It has 62 contact points (electrodes) on the lower surface where the tongue can make a contact. These are distributed in eight rows, which correspond to particular articulatory regions as shown in Fig. 2.1. This palate is connected to the EPG electronic instrumentation which produces running frames showing the points by back squares at which a contact is made.

EPG system also records the speech signal and provides both spectral and wave form representation (Fig. 2.2). All information are time-aligned with the EPG frame representation. Figure 2.2 depicts the production of a plosive in V-V context with occlusion, burst and voice onset indicated therein.

For EPG analysis the informant has to wear the false palate in his/her mouth and hold an electrode in one hand. It takes some time for the informant to get

Fig. 2.1 Articulatory regions of the custom made palate



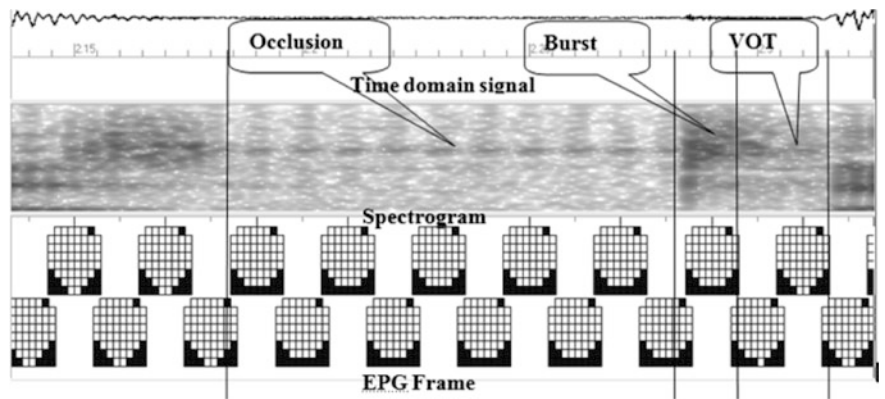


Fig. 2.2 An example of the output from EPG

accustomed. There is no doubt that the system is slightly cumbersome, uncomfortable and highly expensive in the Indian context. However, there being no other more reliable method available for determining the place of contact of tongue and as International Phonetics Association recommends this instrument for assuring standardization of data this has been used for the present study.

2.3 Speech Material

For acoustic analysis of the manner of articulation of consonants, data in VCV context was taken from the speech corpus of DIT, Kolkata [www.cdackolkata.in]. Altogether speech data of 5 speakers (2 male and 3 female) had been considered for this study. All of them are native speakers of Standard Colloquial Bangla. They were selected through a properly constituted selection test. The nativity was assured using three criteria, (1) both the parents must belong to the relevant districts, (2) the informant was reared by their parents, and (3) the education up to the higher secondary level was through Bangla medium. The metadata of the informants is given in Table 2.2. The words used were taken from a phonetically balanced word-set (PBW) of about 600 words. The PBW was prepared such that as far as

Table 2.2 Metadata of the informants

Speaker	Age (years)	Sex
1	25	Male
2	36	Male
3	25	Female
4	36	Female
5	42	Female

Table 2.3 Number of different VCV segments

Type of VCV segments	No. of segments			
	Velar	Retroflex	Dental	Bilabial
Unvoiced unaspirated plosive	256	276	263	242
Unvoiced aspirated plosive	260	151	228	69
Voiced unaspirated plosive	206	49	186	200
Voiced aspirated plosive	117	28	110	65

practicable all necessary CVC combinations used in the language are available and there occurrences are nearly the same. These words were embedded in a neutral carrier sentence for reading and recording.

Furthermore natural words containing the possible CC cluster and the C at three different positions were also added to the list. For the analysis of voiced consonants only those VCV segments were considered in which the occlusion period shows distinct voice bar, and similarly for unvoiced ones only those VCV segments having no voicing in the occlusion region were considered. The number of different VCV segments for different consonants was given in Table 2.3. For the analysis using EPG the unit list consists of a set of nonsense VCV sequences in which V represents the seven Bangla vowels /u/, /o/, /ɔ/, /a/, /æ/, /e/ and /i/ and C represents all the consonants of Bangla. For stops all VC sequences for the 20 stop consonants were added in the list.

It may be seen that for unaspirated voiced and aspirated voiced retroflex plosives the number of data were remarkably low. In Bangla the occurrence of these consonants in VCV context is somewhat rare. This is also observed in case of aspirated unvoiced and aspirated voiced bilabial plosives. For analysis which required EPG the appropriate recording material was read out by one male and female informant, wearing the respective artificial sensor palates. Only one informant of each sex was used because of logistic problem as the palates cannot be made in India. The EPG dynamic contact information was recorded in the PC at a regular interval of 10 ms. The acoustic signal from a microphone supplied with the instrument was recorded at 22,050 sampling frequency 16 bits mono PCM format, synchronized with EPG frames. The reading is limited for a maximum of 20 min duration in a session for the convenience of the informants. Ten repetitions each of the VCV and the VC from the aforesaid list were recorded for each speaker for the determination of place of articulation of consonants.

2.4 Methodologies

The objective analysis of consonantal sounds can have two approaches, one is primarily instrumental and the other uses signal processing techniques to extract acoustical information from digitized audio signals. As has already been mentioned, instrumental approach is not often preferred because it affects the naturalness of free

speech. It is reported that the dynamics of certain parameters like formant frequencies show strong correlation with the place of articulation (Datta and Mukherjee 2011) and is largely used in speech technology applications (Das Mandal 2007; Choudhury 2006; Datta 1988; Datta et al. 1978a, b, 1981; Datta and Ganguly 1981, 1985) for the categorization of plosives/affricates with respect to the place of articulation. It may also be noted that these parameters can be reliably extracted, using appropriate signal processing techniques, from online as well as recorded data and therefore can represent some characteristics of free speech. However the instrument EPG may be used to validate the description of the place of articulation obtained through signal processing techniques. For various reasons, mainly methodological, the categories show overlap significant enough for this approach for resolution of ambiguities. The EPG provides possibly the best way to determine the place of articulation with a reasonable degree of accuracy. Moreover the lacuna regarding the naturalness may be somewhat addressed by allowing the informant to get accustomed with the contraption. It is hoped that with some practice this disadvantage is mitigated to a large extent.

Therefore in the present book the primary emphasis is given on signal processing techniques as far as acoustic parameters are concerned. For the place of articulation the EPG is used.

2.4.1 Acoustic Analysis

The signatures for different manners of production related to voicing and aspiration are available in both the signal domain and spectral domain presentation. Plosives and affricates of Bangla (*spɔɾʃɔbɔɾɲɔ*),¹ like all other languages of the Indo-Aryan family, are traditionally presented in a matrix form where the columns represent manners and rows places of articulation. These show a five-way contrast in terms of both manner and places of articulation. In Table 2.4 these are organized in a 5×5 matrix where the places of articulation are arranged from velar to labial. The categories mentioned in this table are traditional subjective views. Of these the second row presents the affricates the others are plosives/stops. The third row presents the retroflex plosives.

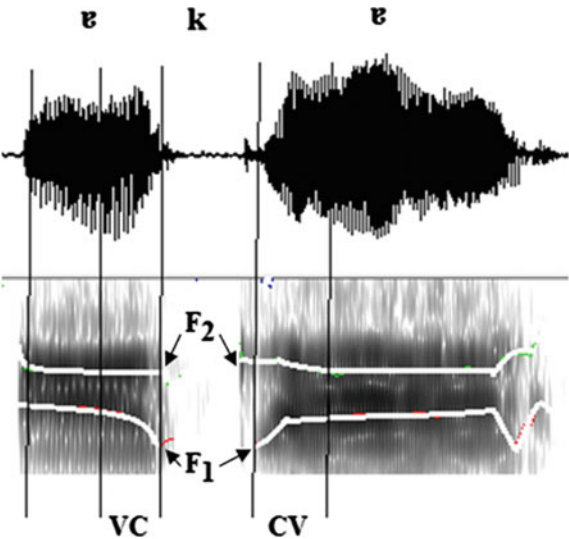
The acoustics of all the consonants are presented below separately for individual plosives, affricates, fricatives, lateral, trill, and taps. An additional section on consonantal murmurs has been included though it is not a separate phoneme, yet it exhibits characteristic acoustic signatures and requires special attention in speech technology applications. These represents voiced occlusion portions of plosives and affricates. The reason behind inclusion of the section is that one does not find in-depth study of acoustics of these portions in literature. It may not be out of order here to mention that the definition of the so called unit ‘phoneme’ is not really a

¹Literally meaning phones produced by touching.

Table 2.4 Traditional matrix representation of Bengali plosives and affricates

	Unaspirated unvoiced	Aspirated unvoiced	Unaspirated voiced	Aspirated voiced	Nasal
Velar	[k]	[k ^h]	[g]	[g ^h]	[ŋ]
Palatal (<i>affricate</i>)	[tʃ]	[tʃ ^h]	[dʒ]	[dʒ ^h]	[ɲ]
Post-alveolar (<i>retroflex</i>)	[ʈ]	[ʈ ^h]	[ɖ]	[ɖ ^h]	[ɳ]
Alveolar	[t]	[t ^h]	[d]	[d ^h]	[n]
Labial	[p]	[p ^h]	[b]	[b ^h]	[m]

Fig. 2.3 Illustration of phoneme boundaries



crisp one. In fact, in real speech it is hard to locate them accurately. Let us consider a simple vowel-plosive-vowel syllable (Fig. 2.3). We can locate the burst and the voice onset time (VOT) as well as the occlusion period quite confidently. But that is not the whole plosive.

The plosive extends into the vowel, so called VC and CV segments, only then it becomes cognitively meaningful. Again the whole voiced, the quasi-periodic region namely from the beginning of CV to the end of the signal in the above figure, is not the vowel alone. If one listens to only this whole quasi periodic signal segment one will hear the CV syllable. For normal listener the difference on including the preceding and the succeeding consonants is not even noticed. Again one has to listen only to the complexity-wise steady part of the signal to hear the pure vowel. Even then the boundaries of the steady part are also fuzzy, at least cognitively. When one studies the acoustics of speech sound, one must give adequate

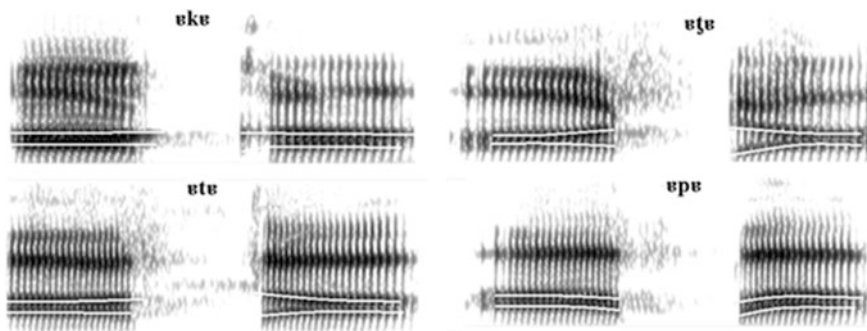


Fig. 2.4 Illustrating transitional cues for place of articulation of plosives

importance to the cognitive domain and be careful about segmentation so that relatively confident zones are selected for parameter extraction.

In the following sections acoustics primarily related to the manners of production will be discussed. It is because the other basis of categorization, namely the place of articulation, does not normally put any distinctly robust imprint on the spectral structure of the different elements of the consonantal segments. The consonantal segments constitute occlusion of the vocal tract, burst, release and segment of aspiration (see Fig. 2.4). Its effect is strongly marked on the quasi-periodic signal representing the adjoining vowel with which the consonants interact strongly. In fact the parameters like amount and direction of formant transition and the time of transition (Datta et al. 1981; Datta and Ganguly 1981), are reported to be distinctively related to the place of articulation. Figure 2.3 illustrates this with a particular VCV syllable. It may be mentioned here that the transition of the second formant provides the best cue (Datta and Mukherjee 2011) for this purpose as this formant represents the backwardness of the tongue hump. One may notice from Fig. 2.4 that the transition of F_1 and F_2 are different for different consonants. While F_1 transition is always downwards towards the burst, since at the point of release tongue always touches the palate and F_1 is related to the height, F_2 movement can be both upwards and downwards. The F_2 transition varies normally according to the place of articulation of the consonant. There is no F_2 transition for [k] in [ɛkɛ] since it has almost the same place of articulation as that of the vowel.

2.4.2 Acoustics of Plosives

Plosive sounds result from the blocking of the vocal tract by the tongue or lips, allowing the air pressure to build up behind the closure, and then by a sudden release of it. This mechanism produces sounds like /p/, /k/ etc. The time for which the tract remains blocked is called occlusion period. In Bangla for each place of articulation there are five manners of production. Of these four manners related to voicing and

aspiration are produced by a unique process of co-ordination between the glottis and the articulators namely the tongue or lips. Nasals are produced by connecting the nasopharynx with the oral cavities by the opening of the velum. The aforesaid co-ordination may be explained with reference to Figs. 2.5, 2.6, 2.7 and 2.8. Each

Fig. 2.5 Mechanism for production of unaspirated unvoiced plosives

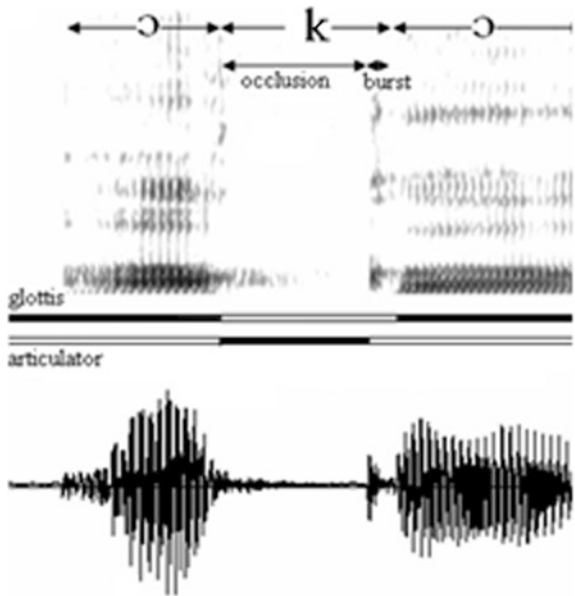


Fig. 2.6 Mechanism for production of aspirated and unvoiced plosive

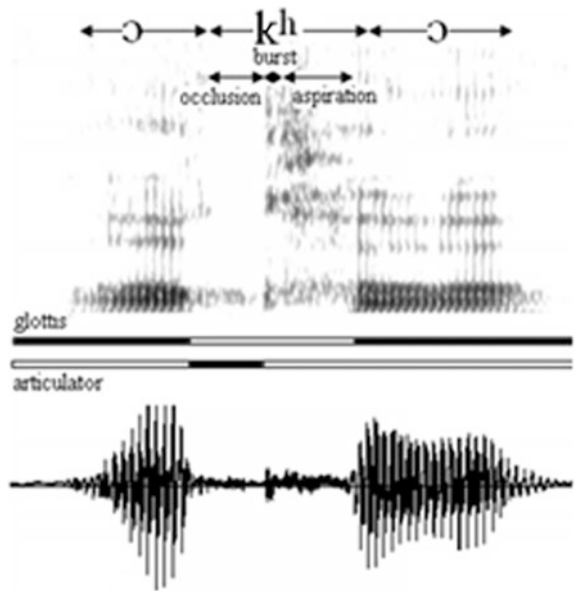


Fig. 2.7 Mechanism for production of unaspirated and voiced plosive

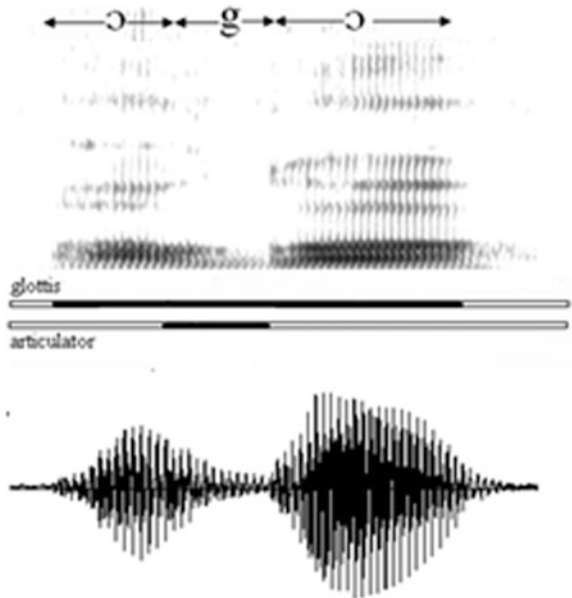


Fig. 2.8 Mechanism for production of aspirated and voiced plosive

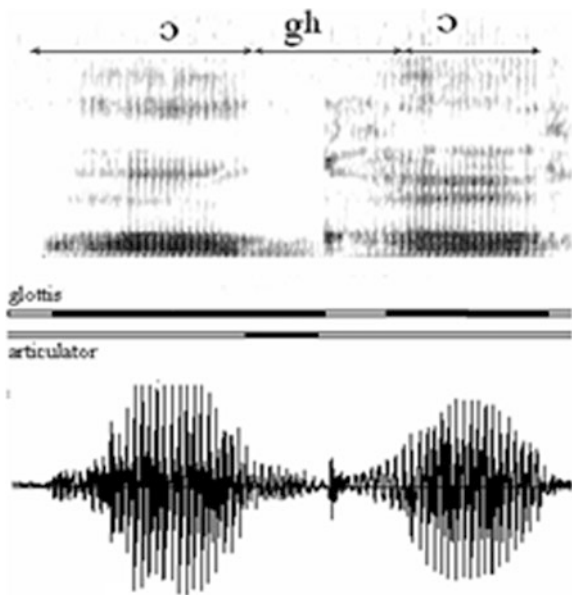


figure has three sections. The uppermost and the lowermost sections show respectively the signal and the corresponding spectrogram. The two bars in the middle section are used to indicate the open (no colour) and the close (black colour) status of the articulator and the glottis. Closed glottis produces the source voice and the closed

articulator suppresses energy in the higher frequencies. The mechanism for the production of different manner of articulation is explained as follows:

- (1) When both glottis and articulators are closed voice bars indicative of the voiced manner of plosive are produced.
- (2) If the articulator remains closed but the glottis is open it indicates a voiceless occlusion.
- (3) When both the articulator and the glottis remain open aspiration is produced because of the turbulence created through the rushing of air through glottis. (One may note that in this particular context the glottis may not be fully open.)
- (4) If along with the condition (1) above the velum is open connecting the nasal cavity, nasal murmurs are produced.
- (5) Closed glottis with open articulator produces normal quasi-periodic sounds.
- (6) If along with (5) velum remains open these give rise to the nasal set of vowels.

Voice Onset Time (VOT) is used in this article to define the time taken to close glottis after the release of closure of the oral tract by the articulator. When the length of VOT is large enough to make it acoustically perceptible, one hears them as the aspirated counterpart of the phonemes even if there is not much aspiration energy. Again if the VOT is small (say <20 ms), even if there is significant aspiration energy the aspiration is not heard.

2.4.3 Acoustics of Affricates

An affricate is a combination of plosive and fricative. In Bangla there are again four types of affricates according to their manner of production, (1) unaspirated unvoiced affricates /tʃ/, (2) aspirated unvoiced affricates /tʃ^h/, (3) unaspirated voiced affricates /dʒ/ and (4) aspirated voiced affricates /dʒ^h/. As we have seen in the case of plosives in last section these manners also reflect unique co-ordination between the opening and closing of the articulators with those of the glottis. However for affricates the important thing to note is that unlike plosive the release is not explosive in nature. The release is rather slow allowing the articulator to make a constriction for some time before the full opening. During the period of constriction air passing through it produces turbulence to cognitively reflect as noise. This together with the occlusion has the cognitive sense of affricates. This part is indicated by the grey in Figs. 2.9, 2.10, 2.11 and 2.12. On the other hand a fricative (as will be discussed in a later section) does not have an occlusion period and the friction is significantly longer. In voiced affricate the occlusion is voiced whereas in unvoiced affricates this part is silence. The spectral structures of the frictional area may be seen from the corresponding fricatives discussed in the next section.

Fig. 2.9 Mechanism for production of unaspirated and unvoiced affricate

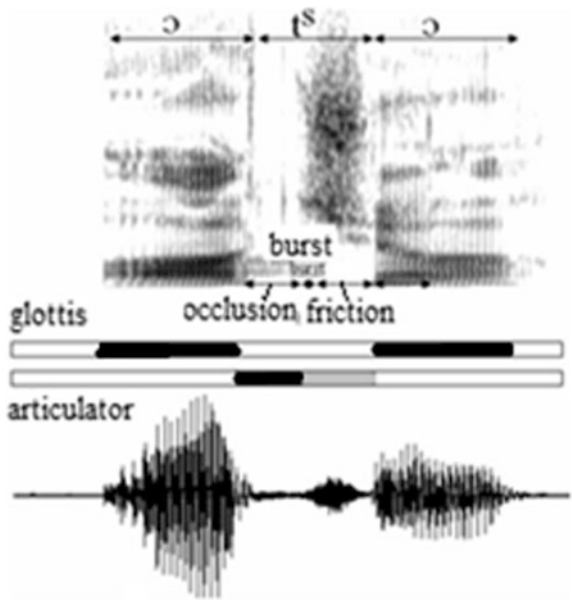


Fig. 2.10 Mechanism for production of aspirated and unvoiced affricate

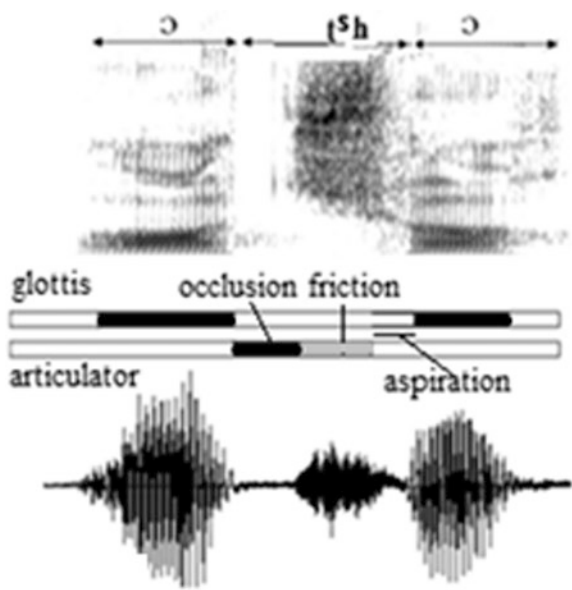


Fig. 2.11 Mechanism for production of unaspirated and voiced affricate

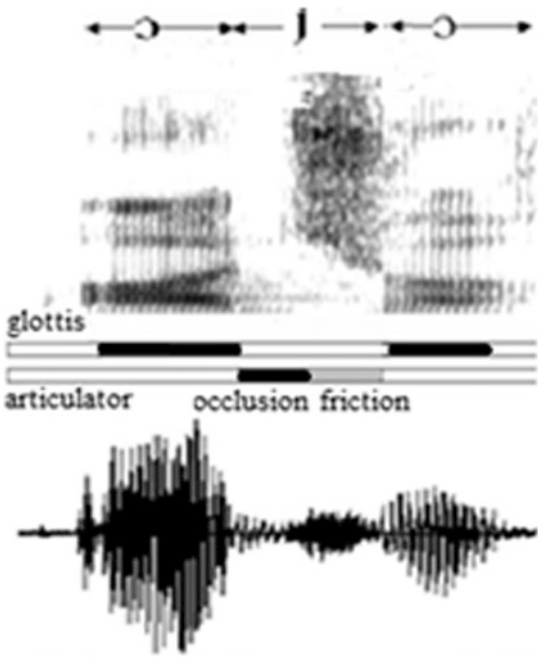
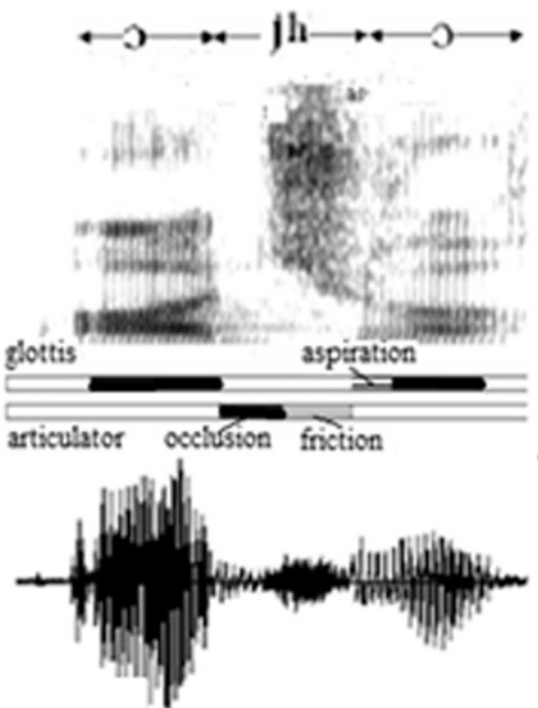


Fig. 2.12 Mechanism for production of aspirated and voiced affricate



2.4.4 Acoustics of Fricatives

Fricative consonants are distinguished from other speech sounds by their manner of production and the fact that they are noisy in character because of their predominant randomness in the spectral structure. While the signal of the voiced speech is quasi-periodic in nature all fricatives are quasi-random in Bangla. Bangla does not have any voiced fricative. The Bangla fricatives are produced by the flow of air through a narrow constriction produced in some region in the vocal tract. The major articulators in Bangla which form the constriction are usually, the tongue blade and the lips. This noise is then filtered by the vocal tract, with the acoustic cavity in front of the constriction contributing the greatest influence in filtering. Figure 2.13 presents the mid-sagittal section of the vocal tract for different place of articulation of the fricatives.

Fricatives are often divided into two groups: sibilants and non-sibilants. Sibilants refer to a cognitive property, a hissing noise, which is the primary way to distinguish between alveolar fricatives and dentals. There may be two different mechanisms for fricative sound generation: (1) an obstacle source, where the turbulence may be generated at a rigid body approximately perpendicular to the airflow, or (2) a wall source where sound is generated along a rigid wall parallel to the flow. For /s, z/ and /š, ž/ the teeth are considered to be the obstacles while the upper lip is an obstacle parallel to the flow for /f, v/, /z, ž/, and /v/ (Shaddle 1990). These are not Bangla fricatives.

Fricative Spectra, according to the theory of speech production and simplified electrical models (Heinz and Stevens 1961), can be characterized by poles and zeroes, which depend on the location of the constriction which is also the source of excitation. Also the resonance frequencies are inversely related to the size of the front cavity. It appears that the back cavity would act as the shunt for the source of noise and therefore their effect on the noise spectra would be revealed as anti-resonances.

There are three fricatives [f], [š] and [s] in Bangla. The basic spectral character of these unvoiced fricatives relate to pure random noise without harmonic structure. The major characteristic spectral features are the resonances and anti resonances. Another important feature is the spectral tilt. Figures 2.14, 2.15 and 2.16 represent

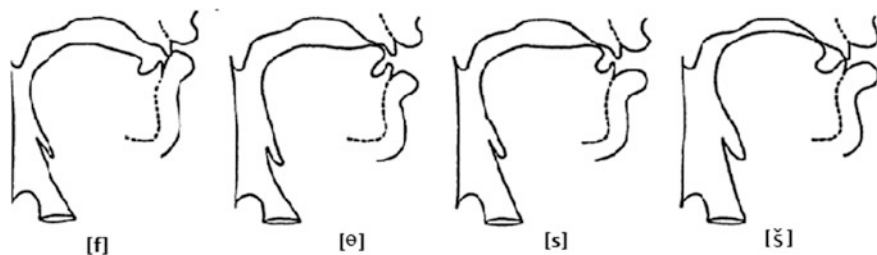


Fig. 2.13 Midsagittal cross-section of a vocal tract constriction (Perkel 1969)

Fig. 2.14 Example of spectra of alveolar fricative of Bangla

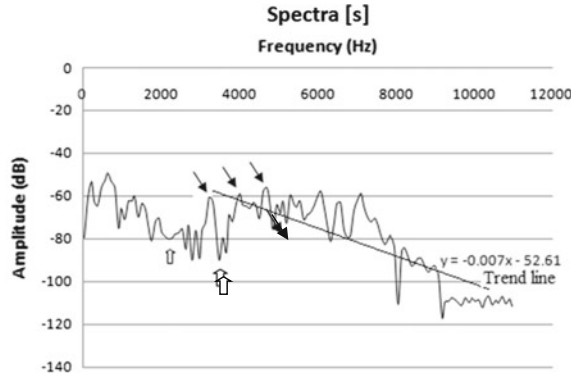


Fig. 2.15 Example of spectra of retroflex fricative of Bangla

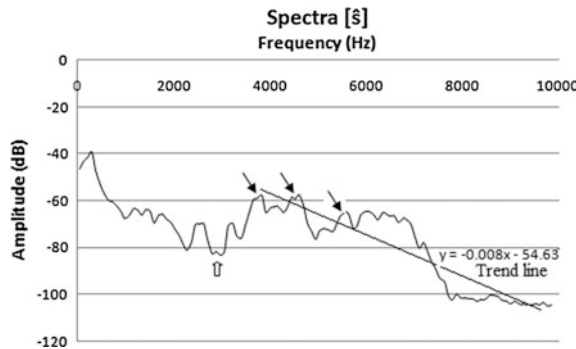
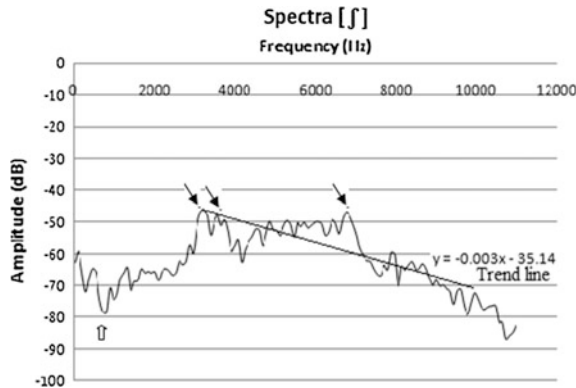


Fig. 2.16 Example of Spectra of palate-alveolar fricative of Bangla



examples of the spectra of the three fricatives in Bangla. There is always a major anti-resonance in the low frequency region of the spectra. This will be referred to as the trough hereafter. In these figures it is indicated by a vertical wide arrow. Only the major peaks after the trough are considered for further description and these are indicated by single inclined arrows. Spectral tilt is considered for characterizing

Table 2.5 Statistics for duration, F₁ and F₂ for Bangla taps and trills

Phone	Statistics	Duration	F ₁	F ₂
r	Mean	51.5	499.5	1423.0
	SD	14.6	150.7	405.7
ɽ	Mean	83.9	353.8	1006.2
	SD	12.1	12.1	305.4
ɽh	Mean	116.2	359.2	1058.8
	SD	13.0	31.3	314.1

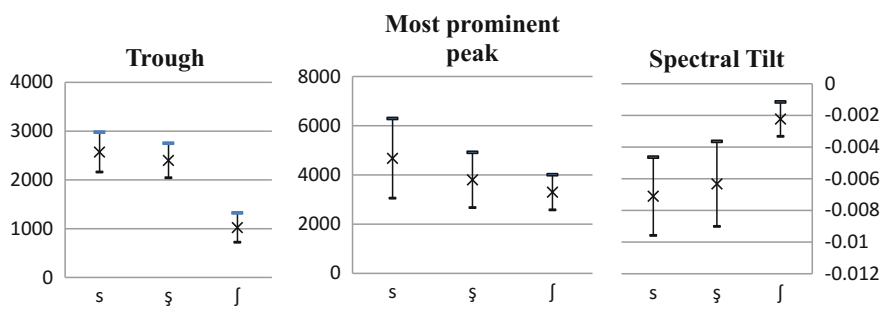


Fig. 2.17 Relative positions of [ʃ], [s] and [ʂ] in three acoustic feature spaces

only the high energy part of the spectrum starting with the first peak point as shown by the trend lines in the figures. The coefficient of ‘x’ in the equation for trend line is taken as the measure of a tilt.

The acoustic features need to be examined are: locations of (1) the trough, (2) the first three peaks, (3) the most prominent peak and (4) the slope of the spectral tilt. The statistics of these acoustic parameters are presented in Table 2.5. Except the tilt all values are in Hertz. For tilt the figures represent the slope. It may be seen from the table that the frequencies of trough, the most prominent peak and the slope of the tilt clearly distinguishes [ʃ] from the other two fricatives. However these parameters are not much distinctive for [s] and [ʂ]. In fact this is vindicated by Fig. 2.17, which represents pictorially the separability of these parameters for the three fricatives. The mean value is indicated by the cross and the lines represent the spreads of the features. The spread equals mean $\pm$ SD. In fact the frequencies of the presented peaks as well as the spectral tilt each makes [ʃ] disjoint from both [s] and [ʂ].

Standard deviations of the frequency of the trough as well as those for peaks for the three fricatives are quite small indicating good stability. There is a small problem though; the peaks do not really represent the mean values of different formant frequencies as they do in vowel spectra. Usually in the spectral data the first, the second or the third peak may appear at any frequency whatsoever.

One could approach to estimate the formants from the probability distribution of the spectral peaks appearing in all the signals of a particular fricative. Figure 2.18 presents these distributions for the three fricatives. It may be seen from these

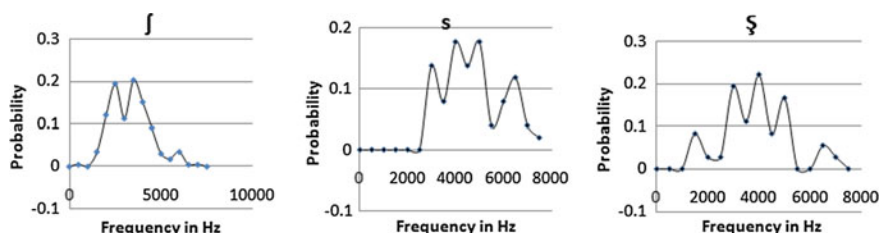


Fig. 2.18 Probability distribution of peaks in spectral energy

diagrams that for [j] two formants at 2500 and 3500 Hz are quite prominent. For [s] the first formant is at 3000 Hz. Three other formants are at 4000, 5000 and 6500 Hz. Similarly for [ʃ] these are respectively at 1500, 3000, 4000 and 5000 Hz. The bins used for calculating the distribution was of 500 Hz width. The aforesaid values represent the mode values not the mean values. One may note that [ʃ] is characterized by a peak appearing at a relatively low frequency of 1500 Hz.

2.4.5 Acoustics of Laterals

The lateral sounds are produced when the mid-coronal closure is complete and the air escapes around one or both sides of the closure. Bangla has only one lateral sound /l/. We shall see later that this could be both bi-lateral and unilateral. For these sounds the air passage is fully blocked medially but at someplace below the dental region an escape route exists some times on both sides (bi-lateral production) and at other times only one side (unilateral production). Because of the fact that the obstruction is at the dental region, the back cavity is large. This should have reflection in the first two formant frequencies. Furthermore the obstruction of the air passage is likely to cause relatively larger reduction of energy in the higher bands though a general loss of energy of the laterals compared to the adjoining vowels is expected. Figure 2.19 represents the spectrogram of a /VLV/ syllable with white line representing the F_0 . The lateral can be easily identified here.

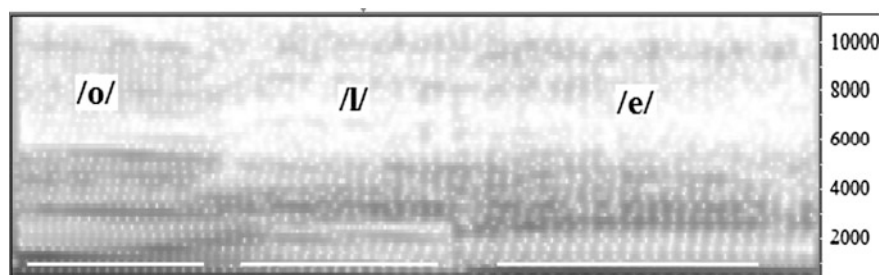


Fig. 2.19 Spectrogram of a /VLV/ syllable

Figure 2.20 gives the spectral sections of the three sections. One could easily see the reduction of the first harmonics.

As mentioned earlier the medial obstruction is likely to create reduction of signal energy. To see whether this happens, it is decided to examine the loss of energy in the low frequency band as well as the high frequency band. For this F0 is determined first and then this value is used as the cut off frequency for band pass filters provided in Cool-Edit Pro (Fig. 2.21)

Figure 2.22 presents the loss of energy on lateralization separately for the low frequency band (F_0) and the high frequency region. There is loss in 96% of the

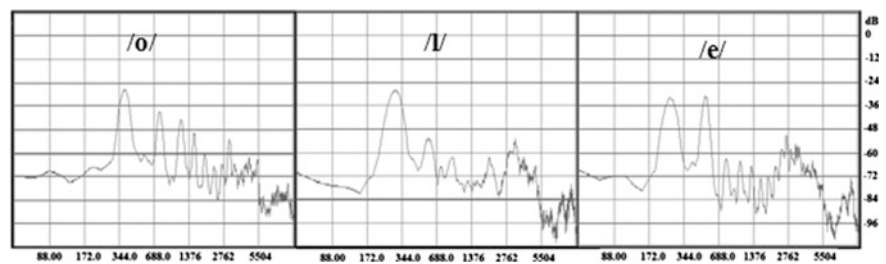


Fig. 2.20 Spectral sections of constituent phonemes

Fig. 2.21 Low pass filter band

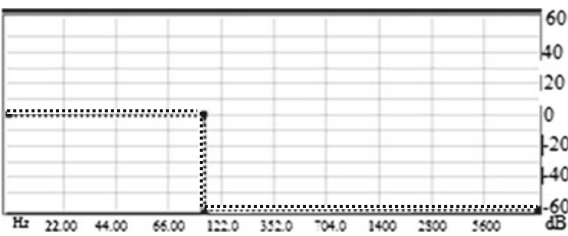
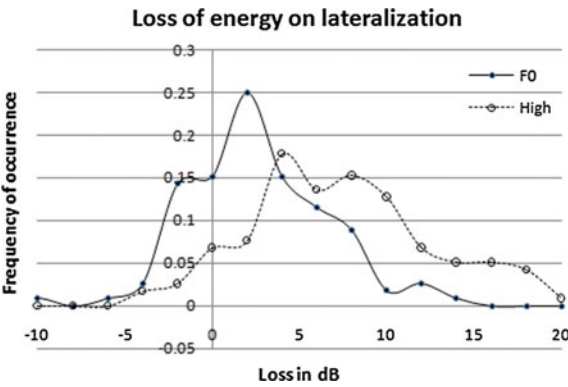


Fig. 2.22 Loss of energy in dB on lateralization



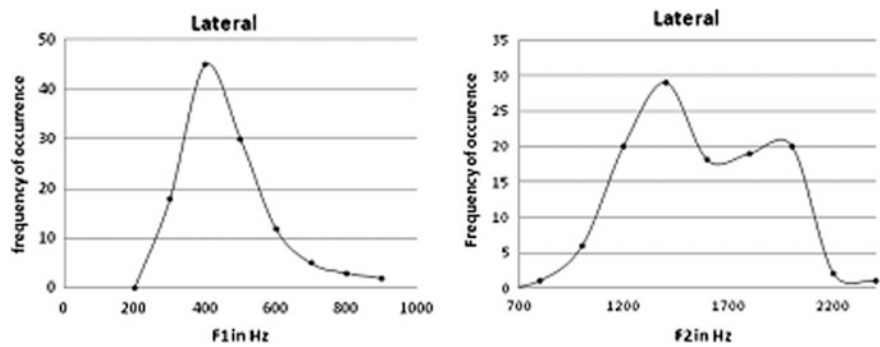
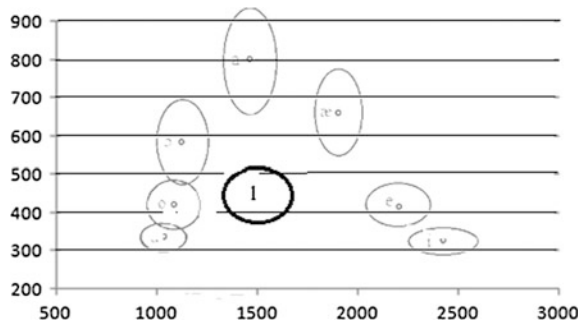


Fig. 2.23 Distribution of laterals in F₁ and F₂ axes separately

Fig. 2.24 /l/ in the F₁-F₂ plane



cases for the laterals in the high frequency band and 81% cases in the low frequency band, the average loss being about 7 dB and 4 dB respectively. Thus there is in general loss of energy in both the bands with that in the F₀ being more prominent.

Figure 2.23 shows the distribution of laterals separately for F₁ and F₂. The mean value of F₁ of the laterals was found to be 421 Hz with the standard deviations of 142 Hz. The distribution of F₂ being far from normal mean value may not be used for representation. One may note that the distribution is bi-modal for the second formant. These two modes are due to the influence of the front-back position of the preceding vowel. The first mode is located at about 1400 Hz and the second mode at about 2300 Hz.

Figure 2.24 shows the position of [l] with reference to Bangla vowels in the F₁-F₂ plane. It may be noticed that the position of the lateral is in the region on inner side of the relatively free space unoccupied by the boomerang shaped vowel distribution. It indicates the possibility of distinguishing the lateral using the first two formant frequencies in case of Bangla.

2.4.6 Acoustics of Trills and Taps

Trills may be described as phones generated by the vibration of some supralaryngeal articulators (tongue-tip, uvula, lips). This vibration is caused by the aerodynamic forces, whereas taps and flaps involve active muscular movements of the tongue. The setting up of the tongue-tip vibration involves muscle contraction, shape and elasticity requirements of the tongue, and a sufficient pressure difference across the constriction. Once trilling is initiated, tongue-tip vibration is maintained as a self-sustaining vibratory system. Pressure builds up behind the lingual constriction until it forces the tongue-tip to open. The air rushes out causing the pressure to drop. The reduction of pressure along the constriction due to the Bernoulli Effect makes the tongue-tip spring back to the contact position. This cycle repeats (Catford 1977; Ladefoged 1967; Spajić et al. 1996; McGowan 1922). For trills the tongue body is more constrained than for the tap. Trills coarticulate less with neighbouring vowels. The acoustic waveform for voiced trills exhibits a low amplitude murmur during the tongue-tip closure.

Tap is produced with a single contraction of the muscles so that one articulator (usually the tongue) is thrown against another articulator. The tongue body is less constrained than in trill. The tap is articulated with a restricted short apico-alveolar closure and more pre-dorsum lowering than other alveolars (Recasens and Pallare 1999).

It is said that Bangla has one trill [r] one unaspirated retroflex tap [ɽ] and one aspirated retroflex tap [ɽʰ]. Figure 2.25 shows examples of spectrograms and wave forms of two instances of [r]. The tap positions are indicated by the arrows. The left one shows a single tap and the right figure shows two taps indicating a trill. The thick horizontal lines above the wave forms indicate the duration of [r]. Similarly Fig. 2.26 shows corresponding examples for [ɽ] and [ɽʰ].

Bangla [r] is traditionally reported to be a trill, which means there would be more than one contact of the tongue with the palate. Similarly [ɽ] and [ɽʰ] is traditionally reported to be taps indicating a single contact. However acoustic examination of a large number of instances of multiple speakers from the aforesaid data base revealed that only 10% of the instances for [r] exhibited a double contact the rests

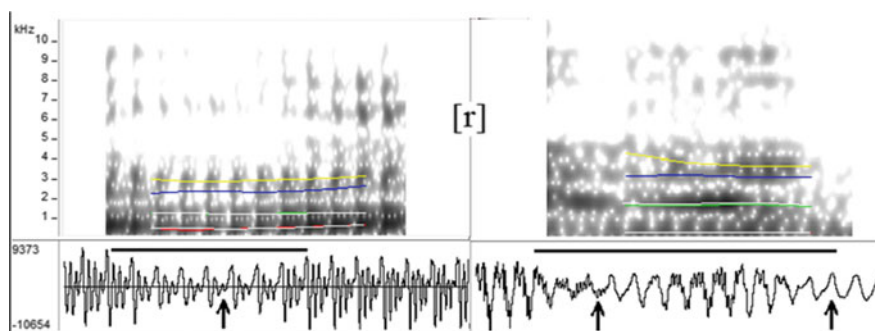


Fig. 2.25 Examples of spectra and wave forms of [r]

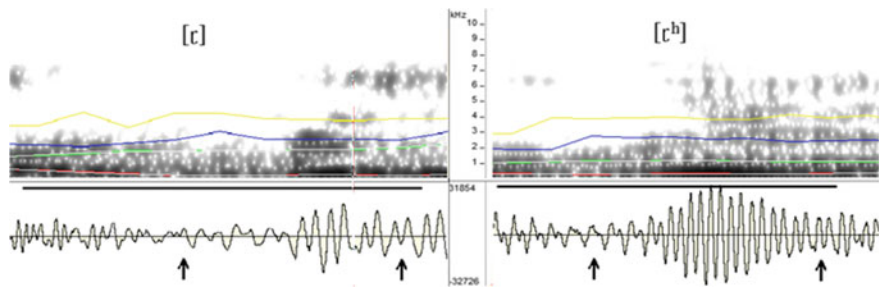


Fig. 2.26 Examples of spectra and wave forms of [ɾ] and [ɾʰ]

only one contact. This indicates that, with Bangla native speakers, [ɾ] is generally a tap. The same examination with [ɾ] revealed that all instances have a single contact indicating it to be a tap. However 88% of [ɾʰ] exhibited double contacts indicating that this may be a trill or a rolled utterance.

Since for all these sounds the oral cavity is divided by the tongue the position of contact is likely to influence the formant frequencies, particularly the F₂. As the data base consists of these sounds in various different vowel contexts the co articulation effect may be quite significant. For locating F₁ and F₂ frequency distribution of the formants has been drawn (Fig. 2.27).

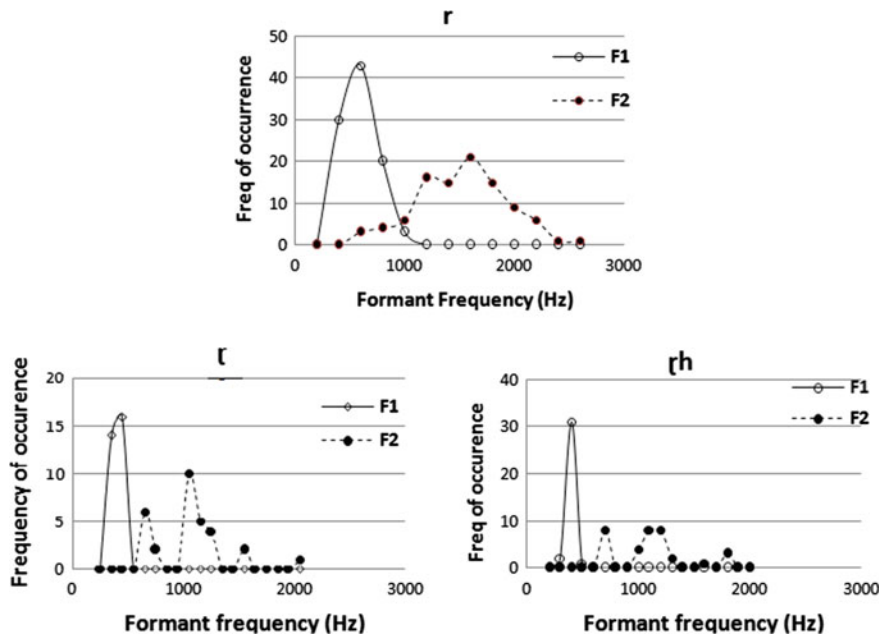


Fig. 2.27 Frequency distribution of F₁ and F₂ for the taps and trills

The distributions for F_1 are unimodal indicating absence of influence of adjoining vowels. Which is expected, as in all cases the tongue touches the palate. Modes observed from Fig. 2.27 for F_1 are 600, 450 and 400 Hz respectively for [r], [ɽ] and [ɽʰ]. The low value for [ɽ] and [ɽʰ] indicate possible lowering of the jaw which may be necessary to accommodate curling of the tongue. The frequency distributions for F_2 exhibit multi modality with major modes at 1200 and 1600 Hz for [r], 650 and 1050 Hz for [ɽ], and 700 and 1100 Hz for [ɽʰ]. The multi-modality indicates contextual influences of adjoining vowels. So does large values of standard deviations.

Table 2.5 presents the mean and standard deviations for duration, F_1 and F_2 for these three consonants. The SDs are generally low for F_1 indicating almost no influence of the adjoining vowels. However the comparatively larger values of standard deviations for F_2 are expected as the distributions are far from normal. A reference to Fig. 2.27 reveals that the distribution of F_2 is multimodal. In fact this puts into question the effectiveness of mean and standard deviation to characterize the variability of this particular parameter. The examination of the detail data revealed strong influence of the vowels on F_2 . The standard deviations of durations for all these consonants are fairly low. Along with that the mean values being quite different, duration may be considered as an inter class distinctive feature for these consonants. Figure 2.28 shows the position of the taps and trills in F_1 - F_2 plane in relation to Bangla vowels represented by gray ovals. The vowel positions are taken from an earlier report (Datta 1988).

Figure 2.29 presents frequency distribution of duration for these consonants. Modes are seen to be well separated. It may be seen that the mode of duration for [r] is around 45 ms, and is distinctively less than those for [ɽ and ɽʰ]. The retroflexing of the tongue is expected to need more time. The modes for the last two are at 105 and 120 ms respectively. The large value for ɽʰ is expected as this trill generally contains two separate contacts.

It needs to be mentioned that there are only few words in Bangla which contains [ɽʰ]. These are /driɽʰɔ/, /muɽʰɔ/ and /ɛʃɛɽʰ/, /gɛɽʰɔ/, /guɽʰɔ/, /mɛɽʰi/, and /ruɽʰɔ/ respectively meaning ‘firm’, ‘dumb’, ‘the name of a month’, ‘thick’, ‘secret’ ‘gum’ and ‘rude’. It is very likely that in common speech [ɽʰ] may not, and even need not, be pronounced in its proper manner for general communication purposes.

Fig. 2.28 Position of Bangla taps and trills wrt vowels

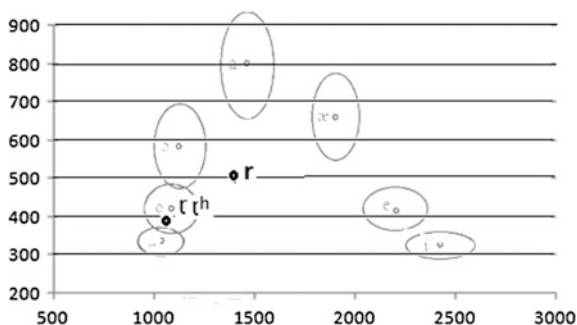
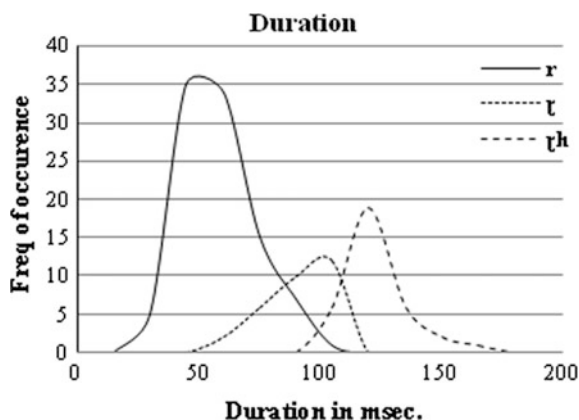


Fig. 2.29 Frequency distribution of duration of [r, ɾ and ɾʰ]



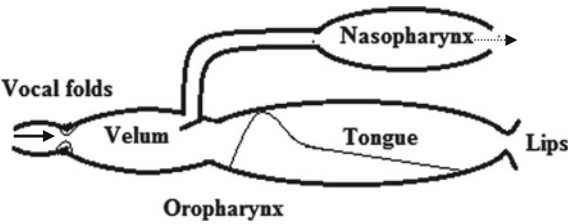
2.4.7 Consonantal Murmurs

Consonantal murmurs are defined here as speech sounds produced by the occlusion of the oral tract when the vocal chords are oscillating. These are not phonemes by themselves but plays an important role in manner distinction and therefore demands investigation. These may be purely oral, e.g. [g, g^h, d, d^h, ɖ, ɖ^h, b, b^h, ɖʒ, and ɖʒ^h] in Bangla or purely nasal, e.g., [n, ɲ, ŋ, m] in Bangla. These murmurs need not be confused with that of murmured vowels, a typical feature in many Indo-Aryan languages (Chap. 1, Sect. 1.5), wherein murmur is associated with a lax glottis introducing aspiration energy into the voice source (Klatt and Klatt 1990; Hillenbrand et al. 1994; Stevens 2000; Ladefoged and Antananzas-Barroso 1985). Though these consonantal murmurs constitute a large portion of consonantal sounds not much attention towards the acoustic structure of these sounds seems to be evident. Major attentions appear to be given to the anticipatory influence of these sounds on the duration of the previous nucleus vowel and to the cognitive aspects of these influences. The objective here is to study only the segmental acoustics of these murmurs.

The articulatory mechanism for the production of these sounds may be understood from the schematic diagram presented in Fig. 2.30. The air stream from lungs causes oscillation of the vocal folds in the closed glottis during the production of the preceding vowel. The murmurs are produced by the occlusion of the vocal tract while the glottis remains closed and therefore voicing source remain active. If during this occlusion the velum also remains closed the oral murmurs are produced. The sound that emerges is not direct but by radiation across the walls of the mouth cavity. However, when the velum is open sound also leaks through the nasopharynx with associated addition of nasal formants.

The closure is likely to cause primarily a severe reduction of the sound energy. One would expect the loss to be primarily in the high frequency region as radiation allows low frequency components to come out albeit with some absorption. As the

Fig. 2.30 Schematic of the articulatory configuration for consonantal murmurs



closure takes place after the production of preceding vowel the air-stream velocity across the glottis is likely to start fading out because the sink is closed. This may result in the lowering of F_0 . Furthermore this closure may also cause a back pressure disturbing the normal free functioning of the vocal folds, the signature of which is likely to be found in the amplitudes of the first two harmonics, H_1 and H_2 , relative to that of F_0 . With these premises in view the study is conducted with about 200 murmurs of which 89 are oral the rest are nasals. A comparative study of the following parameters is done for the consonantal murmurs with respect to the preceding vowels:

- (A) Total energy
- (B) Energy in the low frequency band
- (C) Change in the F_0
- (D) Amplitude of H_1 and H_2 with respect to that of F_0
- (E) The relative amplitudes of H_1 and H_2 .

The VC segments have been manually extracted from the speech corpus of CDAC, Kolkata. The simple affirmative sentences spoken by five male and five female native informants (see Table 1.1 of Chap. 1) has been used for the purpose. The number of oral murmurs and that of nasal murmurs are 89, and 104 respectively. Table 2.6 represents the number of samples for murmurs with different places of articulation.

For the purpose of extracting parameters the signals corresponding to each sentence need to be segmented. This was done manually through an examination of both the signal profile and the corresponding 3D spectrogram.

Figure 2.31 presents the spectrogram and corresponding waveform of a simple affirmative sentence, ‘*ʃe nəbiker nəm kələmbəʃ*’, meaning ‘*that sailor’s name*

Table 2.6 Distribution of samples across murmurs

Murmur	No. of samples
[b]	32
[g]	18
[d]	17
[dz]	21
[m]	26
[n]	87

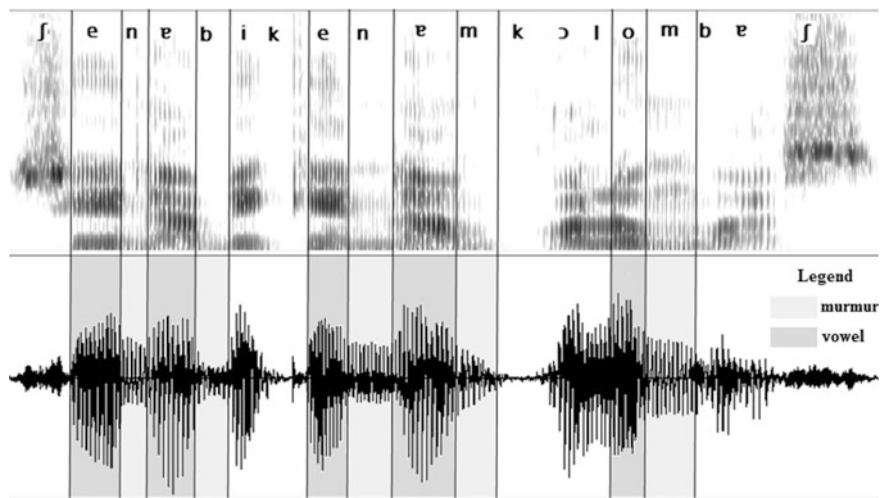
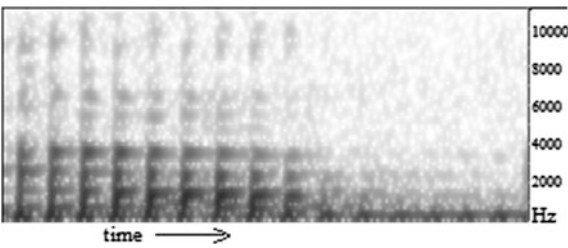


Fig. 2.31 Example illustrating segmentation of the sentence /ʃe nəbɪkɐ nəm kɔləmbʊʃ/

Fig. 2.32 Spectrogram of vowel murmur pair



(is) *Columbus*'. The segmentation of vowels and succeeding consonantal murmurs are marked therein.

The following procedure is adopted generally for the extraction of the parameters. For this segments of signal of length 40 ms are selected on both sides bordering the start of closure of the vocal tract. This closure generally leaves clear identifiable visible signature in both the forms of representation (see Fig. 2.31). However sometimes in both the forms the signature is not equally prominent. For example in a vowel-murmur combination with the vowel [u] or [i] the amplitude profile may not provide a clear separability, in these cases 3D form is helpful. Cool-edit Pro is used for the purpose. Figure 2.32 represents an example of the spectrogram of a typical vowel murmur combination.

Figure 2.33 represents the corresponding spectrogram when the higher energy above F_0 is filtered out using the filter package provided in Cool-Edit Pro. The average loss of total energy due to the closure T_E for the same window of 40 ms is measured using the following simple formulae:

Fig. 2.33 Spectrogram of the same pair after low pass filter

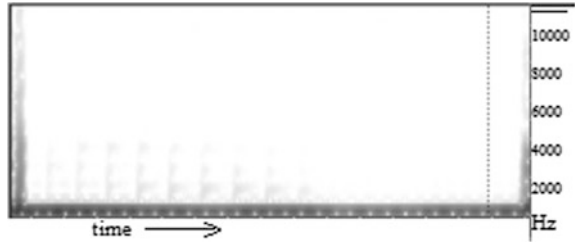
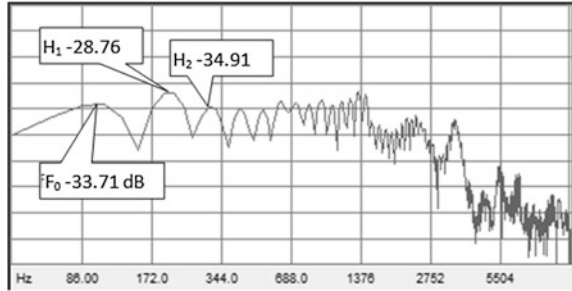


Fig. 2.34 Spectral section showing measurement of amplitudes of F_0 and H_1 and H_2



$T_E = (T_{E_V} - T_{E_M})/T_{E_V}$ and $L_E = (L_{E_V} - L_{E_M})/L_{E_V}$, where T_{E_V} and T_{E_M} represent the average of 40 ms of signal taken respectively just before and just after the closure. Also L_E presents the proportional loss of energy in the low frequency band just covering the fundamental due to the closure and L_{E_V} , L_{E_M} represents the average energy of 40 ms of the filtered signal taken respectively just before and just after the closure.

The spectrum section presented in Fig. 2.34 may be used to describe the measurements of the fundamental frequency F_0 and the amplitudes of F_0 , and of the first two harmonics H_1 and H_2 . The change in F_0 due to the closure is measured in semitone using the formula:

$\Delta F_0 = 12 * \log_2 (^vF_0/^mF_0)$, where $^mF_0$ and $^vF_0$ presents respectively the fundamental frequency of the murmur and that of the preceding vowel.

The sudden closure of the air passage at the mouth cavity is likely to cause accumulation of air coming from the lungs. This may develop a back pressure affecting the vocal fold oscillations which in turn may cause variation in the amplitudes of H_1 and H_2 with respect to that of F_0 . It is necessary to see if there is any such consistent effect. For this, the amplitudes of F_0 , H_1 and H_2 are measured from the average spectrum sections for 40 ms of the vowel signal just before the transition and that of the murmur signal just after the transition. Three differences namely $^A F_0 - ^A H_1$, $^A F_0 - ^A H_2$ and $^A H_1 - ^A H_2$ were computed both for the vowel segment and the murmur segment, where $^A F_0$, $^A H_1$, and $^A H_2$ are amplitudes of F_0 , H_1 , and H_2 respectively.

The occlusion of the vocal tract has been premised here to cause significant loss of energy in the output signal. Figure 2.35 presents the distribution of the proportional losses (PL) of total energy and of low energy due to closure separately for

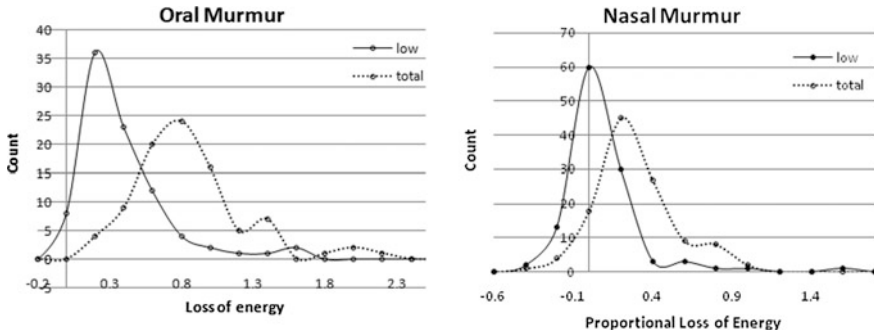


Fig. 2.35 Distribution of proportional loss of energy due to closure

the oral and nasal murmurs. The PL of total energy is seen for all oral murmurs and 96% for the nasal murmurs. Thus as expected, there is generally a loss of total energy for the murmurs. As regards to the PL of energy in the low frequency band just covering the F_0 we see that while there is a loss for all oral murmurs a small but significant portion of about 15% of nasal murmurs do not show any such loss. This may be due to the fact that sound leaks through the nasopharynx with associated introduction of nasal resonances. The mode values are taken to indicate the amount of PL instead of the average values simply because the distributions appear to be far from normal. While with total energy the modal value of PL for oral murmurs is 0.8 dB that for nasal is 0.2 dB. The low value of PL for nasals may be again due to the presence of a parallel path for the sound through the nasopharynx

It is interesting to note that the modal value of the loss in the energy of the fundamental frequency due to closure is 0.8 dB for oral murmurs whereas there is no loss for nasal murmurs. This is because for nasals there is a clear output path through nasopharynx. It seems to indicate that of the total output of energy in the case of oral murmurs the bulk of low energy is through radiation from the walls and that the bulk of the energy directly out of the mouth opening consists of the high frequency band. It must be remembered that the figures given above are proportional loss; actual loss would be much larger as these would be multiplied by the energy of the preceding vowel.

Figure 2.36 presents the change ΔF_0 of fundamental frequency due to the closure of the oropharynx. Again we take the largest mode value instead of average for the reasons stated earlier. The largest mode value of 0 semitones indicates there is generally no change in the F_0 due to closure. However a closer perusal of Fig. 2.36 indicates that about 60% of oral murmurs and 67% of nasal murmurs do indicate a decrease of F_0 . If we consider only samples which indicates a decrease then the average value are 3.8 semitones for oral and 5.4 semitones for nasal murmurs.

This decrease is expected as the air stream velocity across the glottis is expected to fall with the closure of the oropharynx. However under the same expectation on the air stream velocity one would expect the leakage through nasopharynx would work in opposition to the said effect of the closure. One should therefore expect

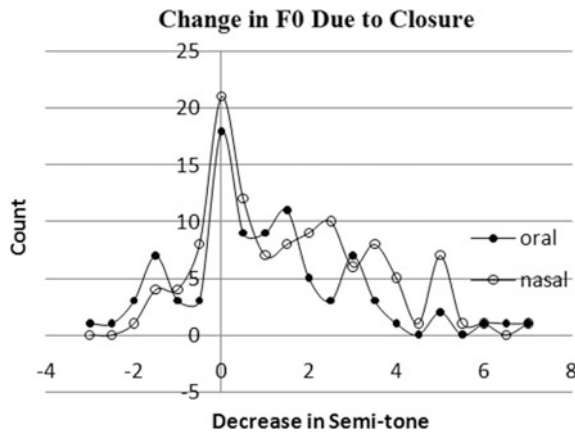


Fig. 2.36 Distribution of ΔF_0 due to closure of the oropharynx

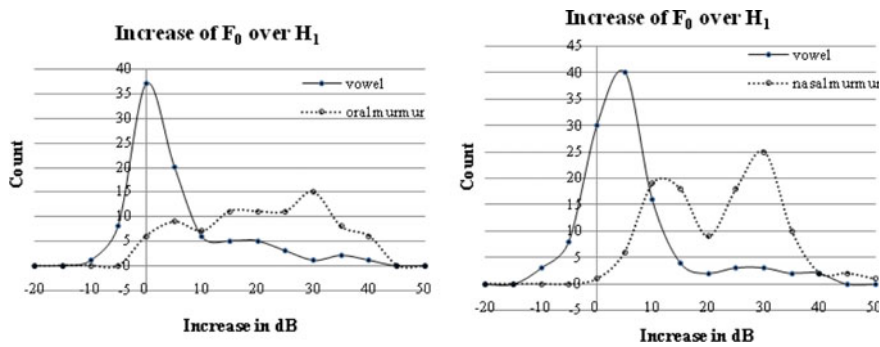


Fig. 2.37 Distribution of samples with the amount of increase of F_0 over H_1

lower value of the decrease for nasals instead of the observed significantly higher values. The main cause of the decrease may however be related to the functioning of the laryngeal oscillators which needs direct observation of the vocal cord movements when the oropharynx is closed and there is a consequent back-pressure generated by accumulated air in the front cavity on the glottal structure.

Figure 2.37 shows the distribution of the decrease of the amplitude of H_1 over that of F_0 for both the vowel segments and the murmur segments. In the case of oral murmurs the modal value of decrease is 0 dB for the vowel and 30 dB for the murmur. However about 48% of the vowels in these pairs showed some decrease while 90% of the murmurs show a decrease of the amplitude. Again for the vowel-nasal murmur pairs modal value of decrease is 5 dB for vowels and that for the murmur is again 30 dB. Results seem to indicate clearly that murmurs cause significant decrease in the amplitude of H_1 over that of the fundamental.

Figure 2.38 shows the distribution of the decrease of the amplitude of H_2 over that of F_0 for both the vowel segments and the murmur segments. In the case of oral

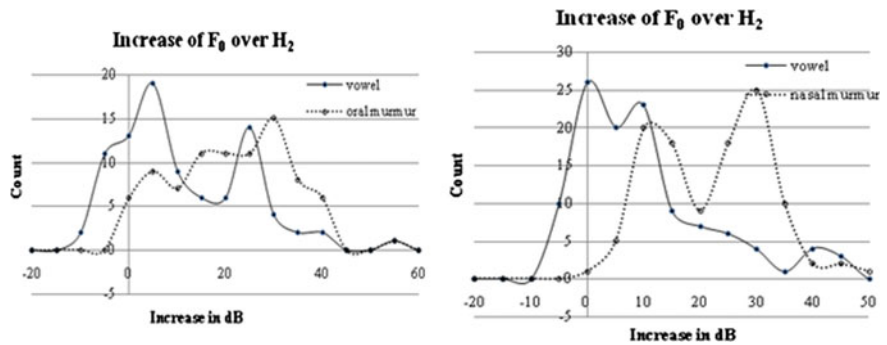


Fig. 2.38 Distribution of samples with the amount of increase of F_0 over H_2

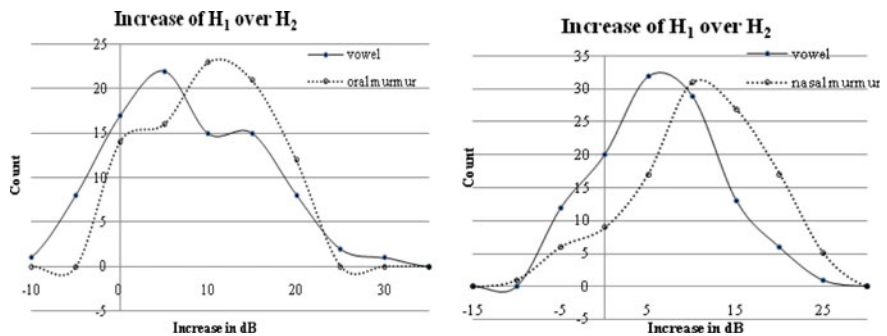


Fig. 2.39 Distribution of samples with the amount of increase of H_1 over H_2

murmurs the modal value of decrease is 0 dB for the vowel though an additional mode of almost equal strength is observed at the modal value of 10 dB. That for the murmur is at 30 dB same as that for H_1 . About 68% of the vowels in these pairs showed some decrease while 98% of the murmurs show a decrease of the amplitude. Again for the vowel-nasal murmur pairs modal value of decrease is 5 dB and that for the murmur is again 30 dB. Results seem to indicate clearly that murmurs cause significant decrease in the amplitude of the second harmonic over that of F_0 .

Figure 2.39 presents the distribution of the increase of the amplitude of H_1 over that of H_2 for both the vowel segment and the murmur segment. In the case of oral murmurs the modal value of increase is 10 dB for the vowel. That for the murmur is at 15 dB. About 71% of the vowels in these pairs showed some increase while 81% of the murmurs show an increase of the amplitude. Again for the vowel-nasal murmur pairs modal value of increase is 5 dB and that for the murmur is again 15 dB. About 71% of the vowels in these pairs showed some increase while 81% of the murmurs show an increase of the amplitude.

An interesting observation from Figs. 2.37 and 2.38 is that nasal murmurs show very distinctive prominent bi-modal distributions. This led to an investigation to see

if they have any relationship with the place of articulation or even with the preceding vowel class. Careful investigation revealed no such correspondences. In fact this bimodality remained even for the two major murmur classes, m and n, as well as for isolated vowel groups. The energy loss due to closure is significant. The loss is most for the higher frequencies, total for oral murmurs but insignificant for nasal murmurs. The loss of energy for F band is negligible. This indicates that in speech the energy output in the very low frequency band is primarily through radiation. The decrease of F_0 frequency has been observed in a large number of cases. A decrease in the average value is 3.8 semitones for oral and 5.4 semitones for nasal murmurs which is unexpected because for nasal a separate path for airflow exists. This put into doubt the conjecture that the decrease is due to the decrease in the velocity of the airflow across vocal folds. The cause of decrease may be related to the functioning of the glottal oscillations due to back pressure generated by the accumulated air in the oral cavity. This gets strong support when we examine the amplitudes of H_1 and H_2 with reference to that of the fundamental. The closure of the vocal tract causes significant decrease in the amplitude of the first and second harmonic with respect to that of F_0 . This indicates the possibility of the effect of closure on the functioning of the vocal folds. One may note here that the closure causes accumulation of air in the back cavity which may cause this.

2.4.8 Summary

While summarizing the aforesaid findings on acoustics of Bangla consonants, it may be pertinent to observe that a comprehensive documentation of acoustics of consonants in any of the Indo-Aryan spoken language is hard to find. Also the author could not find data on acoustic properties of the murmurs associated with voiced plosives and affricates in other language group also. At this moment, therefore, it is not always possible to have an inter-lingual comparison. A crisp summary would therefore be in order here. One important aspect of the plosives and affricates in Bangla, in fact for most of the Indo-Aryan speech sounds, is that they have full repertoire of manner classes. The relation of the acoustics of these manners namely the voicing, the aspiration and the nasalization with the unique process of co-ordination between the glottis, the articulators namely the tongue or lips and the velum are described in detail. It is known that a fully closed glottis produces source voice, while an open glottis, in the present context, produces aspiration and an open velum produces nasality. Different interaction between these three produces the five manners namely, unvoiced-unaspirated, unvoiced-aspirated, voiced-unaspirated, voiced-aspirated and nasal.

For fricatives spectral features are examined with reference to (a) locations of the trough, corresponding to the antiresonance of the cavities prior to the source, (b) first three peaks and the most prominent peak corresponding to the resonances of the cavities, post to the source and (c) to the slope of spectral tilt. The frequencies of the presented peaks as well as the spectral tilt shows [ʃ] to be distinct from both

[s] and [ʃ]. On the other hand [ʃ] is characterized by a peak appearing at a relatively low frequency of 1500 Hz.

Acoustics of the only lateral /l/ is influenced by the medial blockage and lateral escape for the air-stream. There is, in general, compared to the preceding vowel a loss of energy in both the F_0 band and the high energy band with that in the F_0 band being more prominent. The mean value of F_1 and F_2 of the laterals were found to be 421 and 1502 Hz respectively. The respective standard deviations are 142 and 365 Hz. The position of this lateral in the vowel diagram is quite distinct and is in the inner space of the ‘boomerang’ shaped distribution of Bangla vowels. One may note that the distribution is bi-modal for the second formant. These two modes are due to the influence of the front-back position of the preceding vowel.

The acoustic investigations reveal Bangla native speakers [r] is generally a tap. This is in contradiction to the traditional description of [r] as a trill. The same examination with [ɾ] revealed that all instances have a single contact indicating it to be a tap instead of its traditional categorization as a trill. We shall see corroboration of these finding later with EPG analysis. However 88% of [ɾh] exhibited double contacts indicating that this may be a trill or a rolled utterance. The determination of the first two formants helps in fixing their position in the vowel diagram. It is found that while [r] has a distinct position, [ɾ, ɾh] are not separable and their positions are congruent with that of vowel [o].

Consonantal murmurs, a new class introduced for studying their acoustics, are defined here as speech sounds produced by the occlusion of the oral tract when the vocal chords are oscillating. As expected, there is generally a loss of about 80% in total energy and only 30% in the F_0 band. This indicates the most of the output energy in the F_0 band is indirect by radiation through the walls of mouth cavity. For nasal murmurs there is, in general, no loss in F_0 band and only about 30% loss in the total energy. This is also expected as there is an additional pathway for the sound through the nasopharynx. As regards to the effect on F_0 it is found that about 60% of oral murmurs and 67% of nasal murmurs do indicate a decrease of F_0 . The examination of amplitudes of fundamental and the first two harmonics reveal that the amplitudes of the first two harmonics decrease significantly in comparison with those for the preceding vowels.

2.5 EPG Analysis

While acoustics can provide very distinctive description of manners of consonantal categories, it does not provide useful distinctive robust characteristics for the categories related to place of articulation. Apart from subjective categorization, which has been most prevalent traditionally, there are various objective methods of determining the place of articulation. These include cine-fluorography, ultra-sonography, videography, endoscopy and palatography. The cine-fluorography being medically hazardous it is not used now a day. Ultra-sonography has not yet reached the necessary technological level to provide

the needed robustness while endoscopy is very uncomfortable and restricts severely the freedom of the speaker. Thus video capture and palatography remains the main tools in use. The latest development in palatography is the EPG system.

While general principle of using EPG has been briefly described in Sect. 1.2 it is necessary to give some details of experiment conducted for the present study. One issue is the selection of appropriate frame for a particular category of consonant. The other is the methodology used for obtaining a representative frame from a set of frames coming from the repetitions of the same syllable/word. This exercise is necessary for standardization so that the data from different languages are comparable.

Locating the frame or frames for investigating an event must also be well defined. This is necessary as the final data would be a representative of a number of repetitions, and the sequences of each representation should conform to the constituent events in best possible way. An occlusion is easily identified when the frames have the sequence of continuous black cells connect the two sides of the palate showing a complete separation of the front and back halves of the palate (Fig. 1.1 in Sect. 1.2). A closure is indicated when such a frame is preceded by a frame with a discontinuity. A break in this continuity is characteristic of the beginning of a release. The frame characterizing such a break must be seen in the continuity of frames to know whether it is a release or a closure. Contact area is measured by the number of contacts made to the total number of electrodes in the whole EPG frame and is expressed as percentage.

The place of articulation of the friction is determined by the EPG frame where the opening is narrowest. Figure 2.40 shows an example illustrating the procedure for fricatives. Double headed arrow indicates the segment corresponding to the

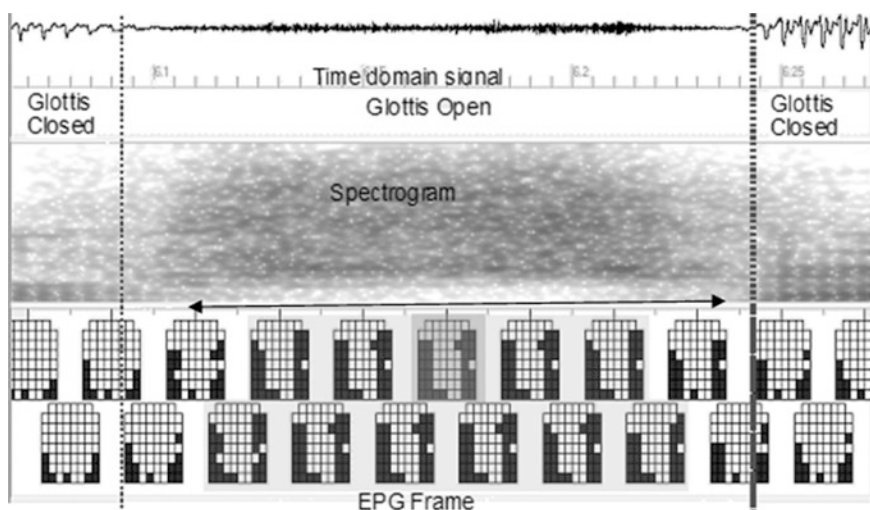


Fig. 2.40 Choice of frame representing the place of articulation for fricatives

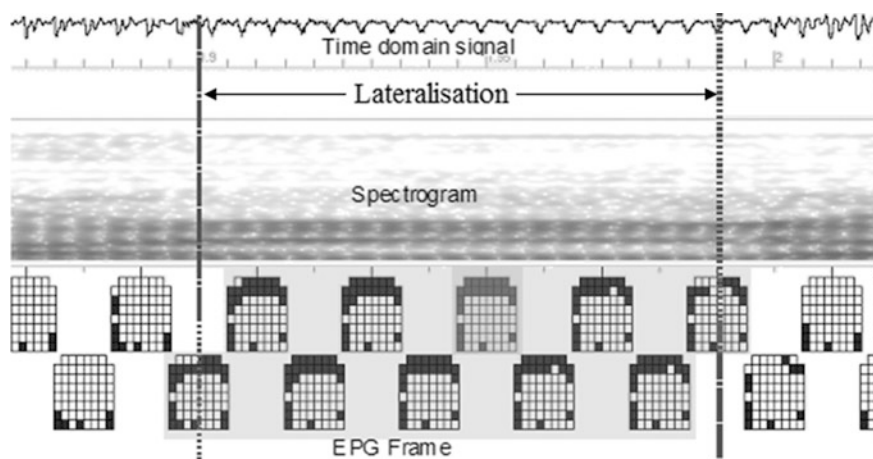


Fig. 2.41 Choice of frame representing the place of articulation for lateral

acoustic event friction. Slightly grayed area represents the frames corresponding to narrowest constriction. The grayer frame at the middle of the friction zone is the selected frame.

Figure 2.41 shows a sequence of EPG frames containing a lateral. The complete medial closure is clear with a large unilateral opening at the right side for this particular instance can also be clearly seen.

For each event like release, closure, constriction etc., related to a phoneme, there are EPG frames the number of which depends on the group being considered, e.g. for individual speaker and a specific context this number is minimum of 10 repetitions. One has to decide on a procedure for getting a frame which shall represent the set of these frames. The issue of registration of the sequence has already been discussed. As indicated in Sect. 1.3 an EPG frame presents an array of small rectangles either black or white. A binary matrix is first obtained by putting '1' for black cell indicating a contact and '0' for white indicating no contact. There are 10 such frames for each event in a sequence, a particular CV or VC context, for each informant. The representative frame is constructed in the following manner. The value of each cell is computed by adding the number of 1's for all these 10 repetitions. Then cumulated values are divided by the number of frames obtained for the speaker for that context. The computed value is used to determine the gray value of the cell where a purely black one indicates that all frames show a contact has been made at that cell. A purely white cell indicates that no contact has been made there. Thus the original binary picture of the frame is converted into a representative gray level picture to accommodate variations in pronunciation.

2.5.1 Plosive/Stop

Figure 2.42 presents one series of EPG frames for VCV utterance where C is an unvoiced unaspirated plosive to exemplify selection of frames to study the place of articulation as plosive and stop for these consonants. The lighter gray shaded frames represents the occlusion period. The deeper gray frames are the terminal ones to be used for determining the place of articulation of plosives and stops.

ক [k]

Figures 2.43 and 2.44 presents the representative EPG frames averaged over the selected frames from ten repetitions, of the release for [k] for each of the seven Bangla vowels respectively for female and male speakers. A general feature that is observed in all such frames is that the contacts begins at the side much earlier going maximally forward to the mid-sagittal position which represents the place of articulation. This indicates a funnel shaped tongue configuration. It can be seen from these that a complete closure is not observed for many vowels due to an inherent restriction of the artificial palate. As mentioned earlier the artificial palate is

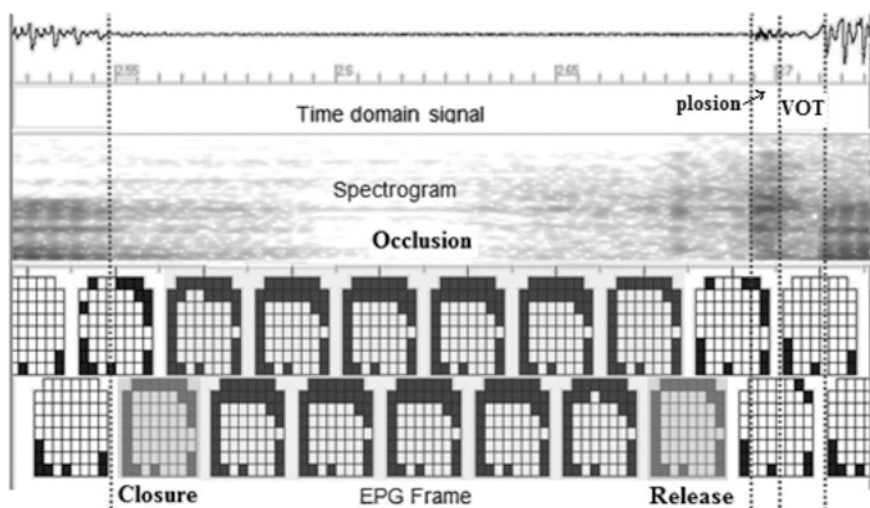


Fig. 2.42 Selection of closure and release frames for plosives

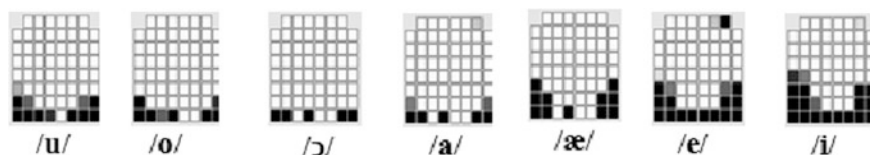


Fig. 2.43 Place of release of [k] for seven Bangla vowels for the female speaker

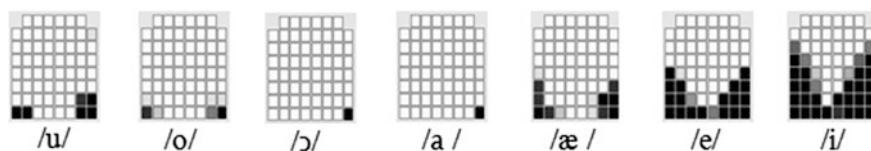


Fig. 2.44 Place of release of [k] for the all Bangla vowels for the male speaker

Fig. 2.45 Overall place of release of consonant [k]

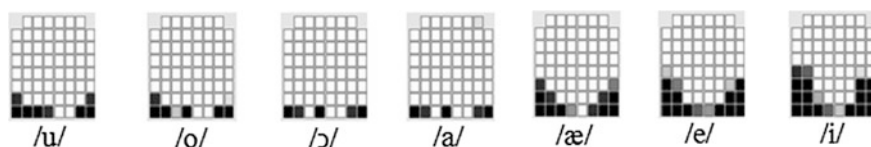
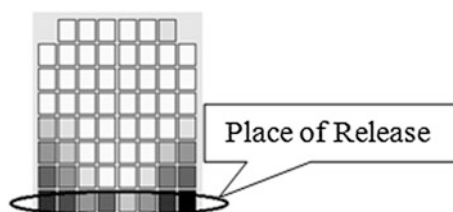


Fig. 2.46 Place of closure of /k/ for the Bangla vowels for the female speaker

made only up to the beginning of the soft palate. This problem will be seen for all manners for this place of articulation. However in the case of front vowels the complete closure is observed. Apparently the tongue hump is pulled forward for front vowels. Also there is a noticeable difference between the male and female informants, the tongue hump seem to be a little bit forwarded for the female speaker.

Figure 2.45 represents the overall place of release for the consonant /k/ with all the frames pooled together. Considering the limitation of the artificial palate mentioned above the figure indicates that the place of release may be taken as the traditional position of velar.

The area of contact of consonant /k/ varies with respect to different vowels. Average duration and energy in the occlusion is found to be 55.76 ms and -42.99 dB respectively. The spectral and signal domain evidences do not show voicing. This together with very low VOT (8.51 ms) makes this plosive unvoiced and unaspirated velar.

Figures 2.46 and 2.47 present EPG frames constructed in the manner described in an earlier paragraph depicting the place of closure of stop [k] for seven Bangla vowels respectively for the female and the male speakers. Here also is a noticeable

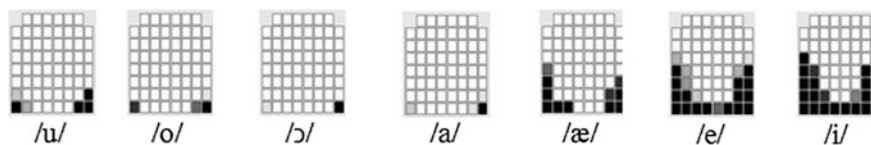
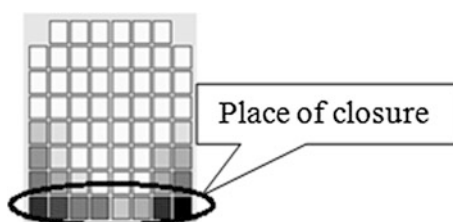


Fig. 2.47 Place of closure of /k/ for the Bangla vowels for the male speaker

Fig. 2.48 Overall place of closure of consonant [k]



difference between the male and female informants, the tongue hump seem to be a little bit forwarded for the female speaker.

Figure 2.48 indicates that the overall place of closure for the stop [k]. Thus [k] is velar both as stop and plosive.

Henceforth the detail EPG frames for individual files shall not be shown in these sections. Only the overall frames computed from all frames separately for the male and female speakers will be presented in these sections. However for the interest of the inquisitive readers these will be provided with appropriate concise comments in the Appendix.

Plosive/Stop ক [k^h]:

Same difference as was seen for [k] between the male and female informants is present for [k^h] also. Figures 2.49 and 2.50 represent the overall places of articulation respectively for release and closure for the plosive [k^h]. Both being almost identical [k^h] is velar both as stop and plosive.

There is not much of a difference in closure configuration between [k] and [k^h] except for a feeble indication that tongue may be a wee bit retracted for the aspirated counterpart.

The area of contact of consonant [k^h] varies with respect to different vowels. Average occlusion duration is found to be 50.23 ms and average occlusion period

Fig. 2.49 Overall place of release of Consonant /k^h/

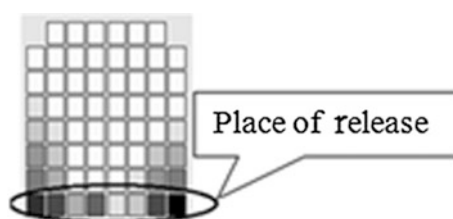


Fig. 2.50 Overall place of closure of consonant [k^h]

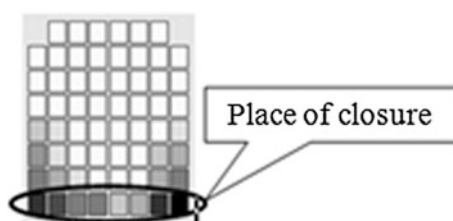


Fig. 2.51 Overall place of release of consonant [g]

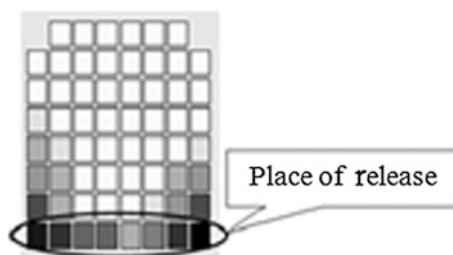
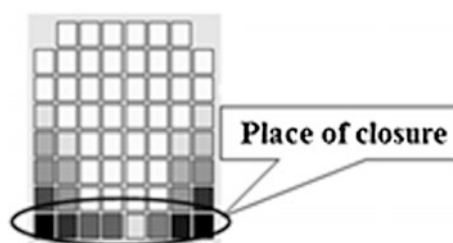


Fig. 2.52 Overall place of closure of consonant [g]



energy is -43.16 dB. The length of the VOT is 46.63 ms which is much larger than 8.51 ms of the unaspirated counterpart [k]. This makes this sound cognitively aspirated. There is no acoustic evidence of voicing during the occlusion. Thus both as a plosive and a stop Bangla [k^h] is unvoiced and aspirated.

Plosive/Stop গ [g]

Same difference as was seen for [k] between the male and female informants is present for [g] also. Figures 2.51 and 2.52 represent the overall places of articulation respectively for release and closure for the plosive [g]. Both being almost identical [g] is velar both as stop and plosive.

Average occlusion duration is found to be 44.42 ms and average occlusion period energy is -31.52 dB. In case of unvoiced counterpart the average energy was -42.99 dB. Thus occlusion region show significant acoustic energy. Also from the acoustic evidence it can be stated that during the occlusion period the glottal vibration do exist. There is also a negligible VOT of 4.47 ms after the burst along with the voiced signal. This duration is non-cognitive as far as aspiration is

concerned. All these taken together shows that Bangla consonant [g] is voiced unaspirated velar.

Plosive/Stop গ [g^h]

Same difference as was seen for [k] between the male and female informants is present for [g^h] also. Figures 2.53 and 2.54 represent the overall places of articulation respectively for release and closure for the plosive [g^h]. Both being almost identical [g^h] is velar both as stop and plosive.

Average occlusion duration is 44.99 ms and average occlusion period energy is -33.01 dB. In case of unvoiced counterpart the average energy was -43.16 dB, which is much smaller. The acoustic evidence shows that during the occlusion period the glottal vibrations do exist. VOT after the burst being long (about 45.89 ms) aspiration is cognizable. Thus Bangla plosive [g^h] is voiced and aspirated.

In general for all manners the average EPG frames for velars, even when examined for each vowel, is almost the same (detail vowel-wise and sex-wise EPG frames are given in Appendix). The effect of the preceding vowels in the closure pattern is generally same, i.e. closure area is pulled forward slightly for front vowels. Similarly for all manners of production the place of articulation is slightly forwarded for the female speaker.

Plosive/Stop ট [t] and ঠ [t^h]

Figures 2.55 and 2.57 represents the overall place of release for the consonant [t] and [t^h] respectively. While the main concentration is in the Post alveolar region significant extension is observed up to the root of the teeth.

Figures 2.56 and 2.58 represents the overall place of closure for the consonant [t] and [t^h] respectively. A comparison of the above two sets of figures reveals noticeable difference. The place of closure for this consonant may still be

Fig. 2.53 Overall place of release of consonant [g^h]

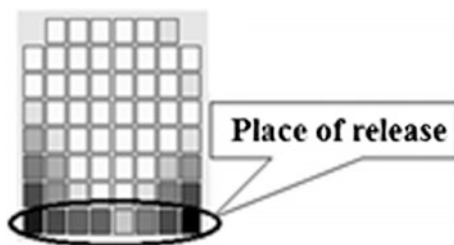


Fig. 2.54 Overall place of closure of consonant [g^h]

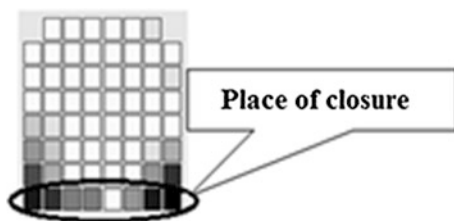


Fig. 2.55 Overall place of release of consonant [t]

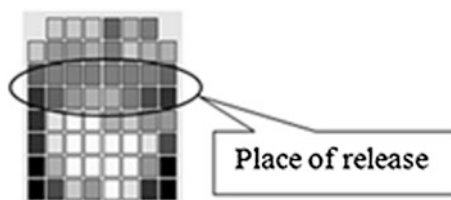


Fig. 2.56 Overall place of closure of consonant [t]

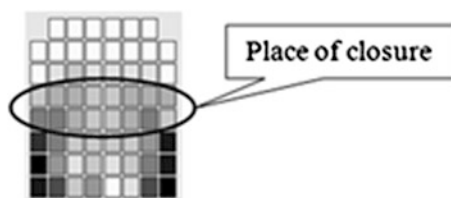


Fig. 2.57 Overall place of release of consonant [t^h]

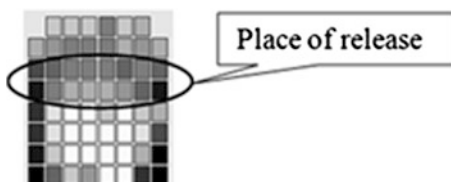
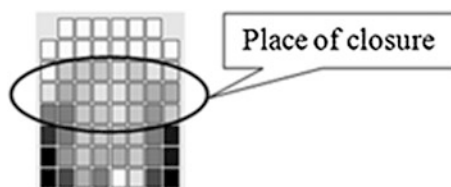


Fig. 2.58 Overall place of closure of consonant [t^h]



considered as post-alveolar but distinctly palatalized. Thus for these plosives the tongue distinctly roles forward for release. Moreover one can notice that more area comes under contact in the case of release.

From the detailed vowel-wise EPGs (see Appendix) it is observed that the place of closure is post alveolar in case of front vowel of both speaker but it is palatal in case of back vowel for both the speaker.

Average occlusion duration for [t] is 59.77 ms and average occlusion period energy is -41.91 dB. No acoustic evidence of voicing during occlusion is found also the length of the VOT (3.06 ms) being insignificant the manner of this plosive is unvoiced unaspirated.

Average occlusion duration for [t^h] is 58.26 ms and average occlusion period energy is -43.22 dB. The significant length 44.05 ms of the VOT and absence of

Fig. 2.59 Overall place of release of consonant [d]

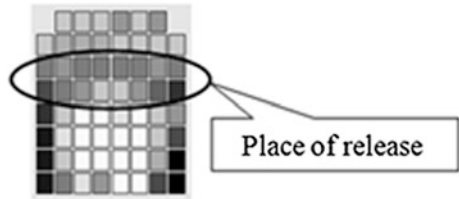


Fig. 2.60 Overall place of closure of consonant [d]

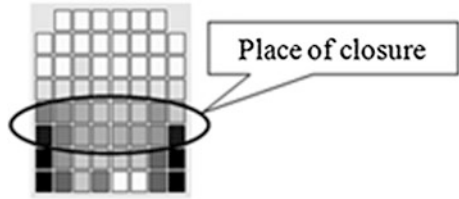


Fig. 2.61 Overall place of release of consonant [d^h]

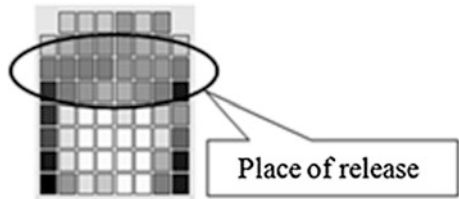
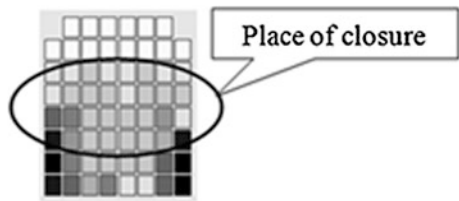


Fig. 2.62 Overall place of closure of consonant [d^h]



the acoustic evidence of glottal vibration during the occlusion makes Bangla consonant [t^h] as unvoiced aspirated post-alveolar as a plosive as well as a stop.

Plosive/Stop ढ [d] ढ

Figures 2.59 and 2.61 represent the overall place of release for the consonant [d] and [d^h]. From this figure the place of these plosives is seen primarily to be post alveolar however one notices like in the case of its unvoiced counterpart some extension up to the root of teeth.

A perusal of Figs. 2.60 and 2.62 shows that the place of closure is palatal. From the detailed vowel-wise EPGs (see Appendix) it is observed that the place of closure is post alveolar in case of front vowel of both speakers but it is palatal in case of back vowel for both of them.

Average occlusion duration for [d] is 59.57 ms and average occlusion period energy is -30.99 dB. From the acoustic evidence it can be stated that during the occlusion period the glottal vibration do exist. The average length of the VOT being 3.31 ms the plosive is unaspirated. So it can be concluded that the Bangla consonant [d] is voiced unaspirated retroflex plosive.

Average occlusion duration of [d^h] is 54.33 ms and average occlusion period energy is -28.67 dB. In case of unvoiced counterpart the average energy is -43.22 dB, which is much smaller. From the acoustic evidence it can be stated that during the occlusion period the glottal vibration do exist. There is also an aspiration of 48.7 ms after the burst along with the voiced signal. The manner of articulation is known to be retroflex. So it can be concluded that the Bangla consonant [d^h] is voiced aspirated retroflex plosive/stop.

It is obvious that EPG cannot provide direct evidence of retroflexion. Therefore one has to depend on the subjective judgment of the speaker or an appropriate video-graphic arrangement for this. However if one carefully peruses and compares the overall EPG frames of all the plosive/stop consonant of Bangla one may notice significantly large increase in the area of contact for the retroflex ones, irrespective of the manners of production. This may be due probably to the fact that retroflexion induces a larger and firmer contact with the palate. This then can be taken as the signature for retroflexion in EPG. The other signature appears to be that the release place is distinctly forwarded from that of closure indicating a forward movement of the tongue.

Plosive/Stop ত, থ, দ, ধ [t, t^h, d, d^h]

Figures 2.63, 2.64, 2.65, 2.66, 2.67, 2.68, 2.69 and 2.70 represent the overall place of release and closures for the consonant [t, t^h, d, d^h].

It may be observed from the above figures that the places of release of these plosives are at the region of root of teeth. Since the first sensor row of the artificial

Fig. 2.63 Overall place of release of consonant [t]

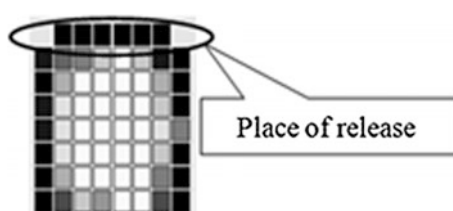


Fig. 2.64 Overall place of closure of consonant [t]

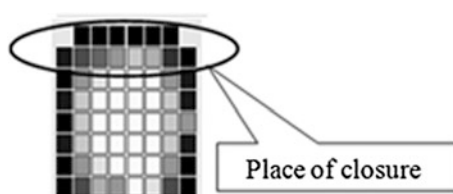


Fig. 2.65 Overall place of release of consonant [t^h]

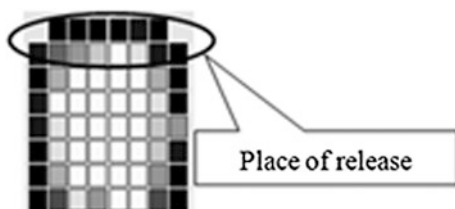


Fig. 2.66 Overall place of closure of consonant [t^h]

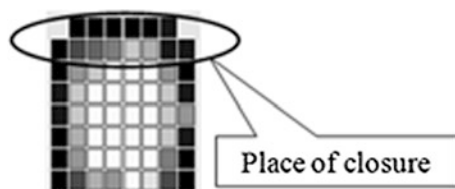


Fig. 2.67 Overall place of release of consonant [d]

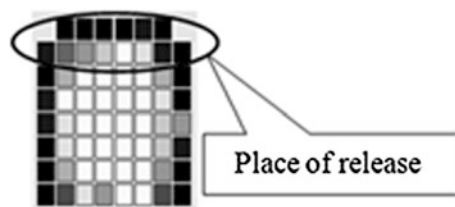


Fig. 2.68 Overall place of closure of consonant [d]

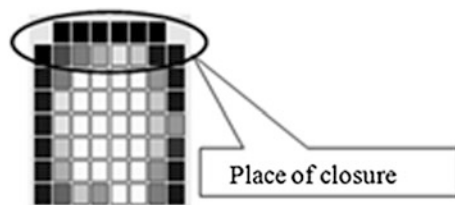


Fig. 2.69 Overall place of release of consonant [d^h]

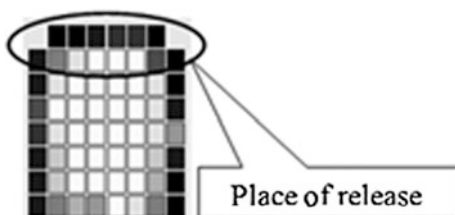
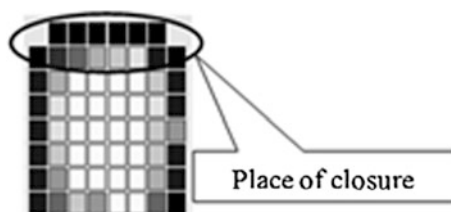


Fig. 2.70 Overall place of closure of consonant [d^h]



palate are only in the root of teeth region (there being no contact for teeth) and contact is always present in this region for all of the vowels this may possibly be considered as dental as it is traditionally believed.

From the above figures it is observed that the place of closure is somewhat retracted but still may be called dental.

Average occlusion duration is 64.93 ms and average occlusion period energy is -43.76 dB. There is no evidence of voicing. The length of the VOT is 3.94 ms which is negligible. Thus the Bangla consonant [t] is unvoiced unaspirated **dental** plosive/stop.

Average occlusion duration for [t^h] is 49.32 ms and average occlusion period energy is -42.56 dB. There is no acoustic evidence of voicing. The length of the VOT is 28.02 ms which makes aspiration cognizable. Thus the Bangla consonant [t^h] is unvoiced aspirated **dental** plosive.

Average occlusion duration for [d] is 48.86 ms and average occlusion period energy is -31.06 dB. From the acoustic evidence it can be stated that during the occlusion period the glottal vibration do exist. There is also an aspiration after the burst along with the voiced signal. The average length of the aspiration part is 3.03 ms and is insignificant. Thus the Bangla consonant [d] is voiced unaspirated dental plosive/stop.

Average occlusion duration for [d^h] is 42.05 ms and average occlusion period energy is -30.64 dB. From the acoustic evidence it can be stated that during the occlusion period the glottal vibration do exist. There is also 48.35 ms of VOT after the burst along with the voiced signal which makes aspiration cognizable. Thus the Bangla consonant [d^h] is voiced aspirated dental plosive.

The limitation in artificial palate to record properly the closure in the velar region has already been mentioned. One may also note here that because the artificial palate does not cover the teeth and only ends at the root of the teeth one is not sure whether some plosive/stops are purely dental.

The vowel-wise EPGs (presented in the Appendix) reveal that both the place of release and closure for both male and female speakers for front vowels the place of contact is forwarded. Though this is not large enough to merit particular attention in general, it is found to be significant in the following situation for retroflex plosives/stops:

- (a) [ʈ, ʈʰ, and ɖ] are dental with male speaker for release in front vowels
- (b) [ʈ, and ʈʰ] are palatal/mediopalatal with both speakers in closure of back vowels

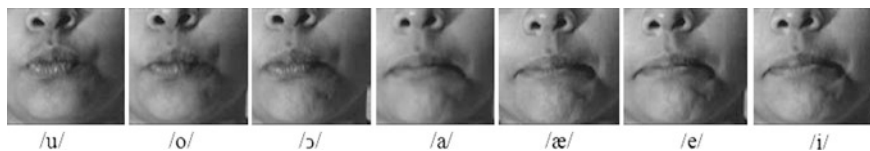


Fig. 2.71 The release frames of lips for [p]



Fig. 2.72 The release frames of lips for [pʰ]

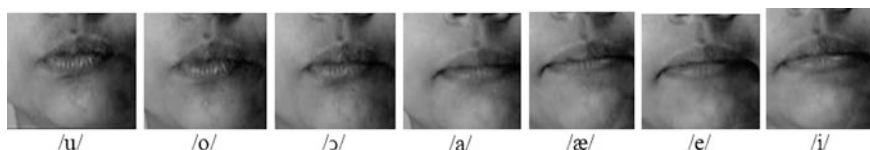


Fig. 2.73 The release frames of lips for [b]



Fig. 2.74 The release frames of lips for [bʰ]

(c) [dʰ] is palatal/mediopalatal with male speaker in closure of back vowels

Plosive/Stop প, ফ, ব, ভ [p, pʰ, b, bʰ]

Traditionally [p, pʰ, b and bʰ] are known as bilabials and as expected the EPG did not show any closure in the entire artificial palate region. The video of lips movements of the informant is recorded during the production of this sound and from the video evidence both the release and closure are found in the lips region. Figures 2.71, 2.72, 2.73 and 2.74 represent the lips release frame for these plosives. These indicate that these plosives are bilabial.

The varied degree of rounding of the lips corresponding to the vowels may be observed from the squeezing of the lips in the pictures.

Table 2.7 Summary of occlusion parameters VOT

	Occlusion		VOT(ms)		Occlusion		VOT(ms)
	Time (ms)	Energy(dB)			Time(ms)	Energy(dB)	
k	55.76	-42.99	8.51	g	44.42	-31.52	4.47
kh	50.23	-43.41	46.63	gh	44.99	-33.01	45.89
t	59.77	-41.91	3.06	ḡ	59.57	-30.99	3.31
ṭh	58.26	-43.22	44.05	ḡh	54.33	-28.67	48.7
t	64.93	-43.76	3.94	d	48.86	-31.06	3.03
th	49.23	-42.56	28.02	dh	42.05	-30.64	48.35
p	70.97	-43.32	3.58	b	51.72	-30.64	2.76
ph	47.51	-41.23	32.11	bh	49.59	-29.94	52.51
Average	57.08	-42.8	NA		49.44	-30.81	NA

Average occlusion duration for [p] is 70.49 ms and average occlusion period energy is -43.32 dB. No acoustic evidence of voicing was found. The length of the VOT, 3.58 ms, is insignificant. Thus Bangla [p] is unvoiced unaspirated bilabial.

Average occlusion duration for [p^h] is 47.41 ms and average occlusion period energy is -41.23 dB. The length of the VOT is 32.11 ms which indicates cognitive level of aspiration. There is no acoustic evidence of glottal vibration during the occlusion period indicating this Bangla consonant to be unvoiced aspirated bilabial plosive.

Average occlusion duration for [b] is 51.72 ms and average occlusion period energy is -30.64 dB. There is also an aspiration after the burst along with the voiced signal. However the VOT of 2.76 ms is insignificant. Along with these, glottal vibration during the occlusion period indicate that the Bangla consonant [b] is voiced unaspirated bilabial.

Average occlusion duration for [b^h] is 49.59 ms and average occlusion period energy is -29.94 dB. The average length of the VOT is 52.51 ms is significant. Along with this glottal vibration during the occlusion period indicate that the Bangla consonant [b^h] is voiced aspirated bilabial.

Table 2.7 gives the summary of parameter values for the occlusion period and VOT for all plosives. The occlusion time is 14% less for voiced counterpart for all articulatory positions. This difference is known to be cognizably significant. The acoustic energy during occlusion is significantly less (12 dB approximately) for unvoiced counterpart. VOT is negligible for unaspirated counterpart and cognizably significant (except for [t^h] and [p^h]). These figures show that manner based classification of Bangla plosives are phonetically significant.

2.5.2 Fricative

The place of articulation of the friction can be determined from the EPG frame where the opening is narrowest (see Fig. 2.41). The method of selection of appropriate frame is described earlier in Sect. 1.5.

Fricative ऋ [ɹ]

Figure 2.75 presents the overall place of constriction for the fricative [ɹ].

It may be seen from the above figure that the overall place of constriction for the consonant [ɹ] is post-alveolar. There being no acoustic evidence of voicing it may be said that [ɹ] is Unvoiced Post alveolar Fricative

There has been no noticeable overall significant difference between male and female informants. However it has been noticed that in case of male speaker the place of constriction is slightly palatalized for back vowels.

Fricative ण [ɳ]

Figure 2.76 presents the overall place of constriction for the fricative [ɳ].

The place of constriction of the consonant [ɳ] is primarily alveolar though extends up forward a little. There being no acoustic evidence of voicing this consonant is unvoiced alveolar fricative. For the informants used the constriction appears to be on the right side of the mouth.

Examination of detail EPGs reveal that in case of front vowel the constriction area is little bit forwarded. In the case of male speaker the constriction area spread up to the root of the teeth region.

Fricative ष [ʃ]

Figure 2.77 represents the overall place of friction for the consonant [ʃ] which is palatal. The retroflexion is not revealed in EPG; however this fricative is traditionally believed to be retroflexed.

Here being no acoustical evidence of voicing [ʃ] may be regarded as unvoiced retroflexed palatal fricative.

Fig. 2.75 Overall place of constriction of consonant [ɹ]

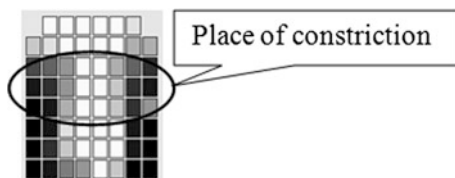


Fig. 2.76 Overall place of constriction of consonant [ɳ]

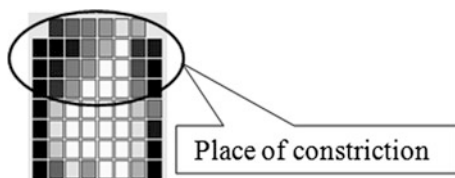


Fig. 2.77 Place of constriction of [ʃ] for pooled data

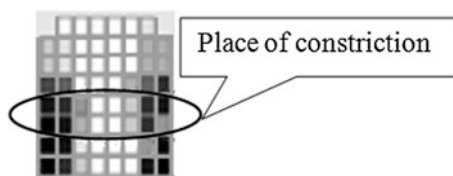
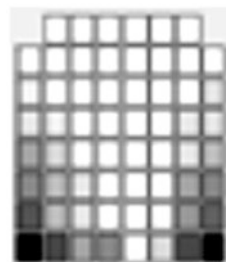


Fig. 2.78 Overall EPG frame for [h]



There is a consistence EPG evidence of the place of articulation of all fricatives in conjunction with front vowels, particularly for the male speaker, to be forwarded.

Fricative ष [h]

Figure 2.78 represents over all EPG frame contact position for all seven aspirated vowels due to the constriction in association with the phone [h].

From the above figures it may be observed that there is an indication of constriction appearing at the lower end of artificial palate. This strongly suggests a constriction below the velar region. The traditional literatures also present [h] as a glottal fricative.

However a recent study of this sound in Bangla (Datta 2014) revealed that in this dialect [h] does not reveal itself as a fricative but causes the succeeding vowel murmured (see Sect. 1.5 of Chap. 1). In fact due to its influence a set of seven murmured vowels has been reported for Bangla.

2.5.3 Affricates

In the case of affricates the place of release of the occlusion is determined from the EPG frame just before the opening of the closure. The place of articulation of the friction is determined by the EPG frame where the opening is narrowest. Manner of articulation of the affricates is determined from the acoustic study of the above-mentioned segments. The average occlusion duration and VOT duration along with the average energy at occlusion period is presented.

Affricate ढ [ɟ]

An affricate is a complex sound in the sense that it contains two consecutive articulatory events namely plosion and friction, physically a sharp short release

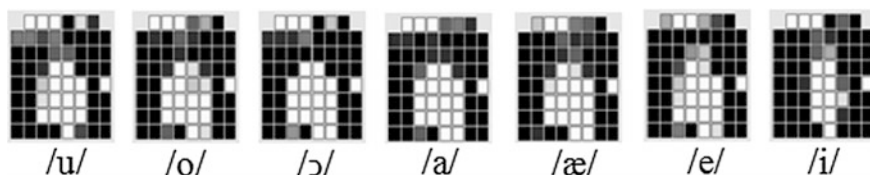


Fig. 2.79 Place of release of [tʃ] for seven Bangla vowels of female speaker

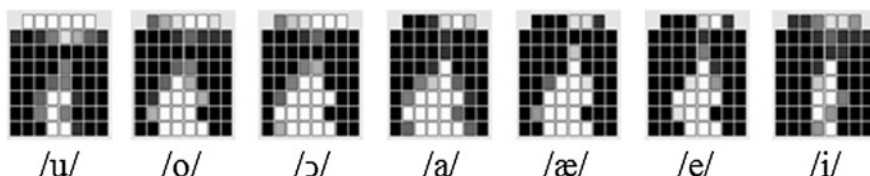
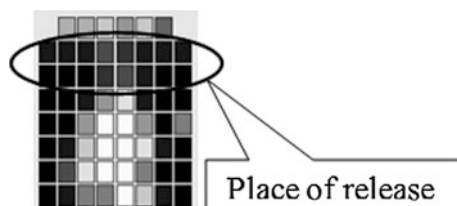


Fig. 2.80 Place of release of [tʃ] for seven Bangla vowels of male speaker

Fig. 2.81 Overall place of release of [tʃ]



followed by a constriction before fully opening for the vowel. During the whole period the tongue may roll. Thus there may be two different places where tongue is engaged with the palate. This calls for citing both of them as places of articulation. However when this consonant is considered as the stop one may take the place of closure as the place of articulation. Similarly, when this is at the beginning of the vowel the place of constriction may be taken as the place of articulation. This convention is followed here.

Figures 2.79 and 2.80 represents the place of release of [tʃ] for seven Bangla vowels spoken respectively by the female and male speakers. Though there is no mentionable difference in the place of articulation the male speaker show greater amount tongue-palate contact.

Figure 2.81 represents the overall place of release.

From the above figures it is observed that the place of release is alveolar for the back vowel in both male and female speaker. In the case of front vowels the tongue tip is little bit forwarded during the release.

Figures 2.82 and 2.83 represents the place of constriction, for generating friction, in [tʃ] for seven Bangla vowels spoken respectively by the female and male speakers.

Figure 2.84 represents the overall place of constriction.

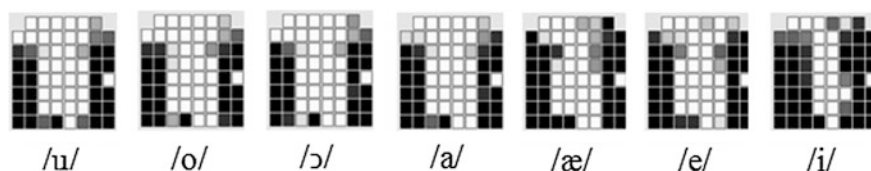


Fig. 2.82 Place of constriction of [tʃ] for seven Bangla vowels of female speaker

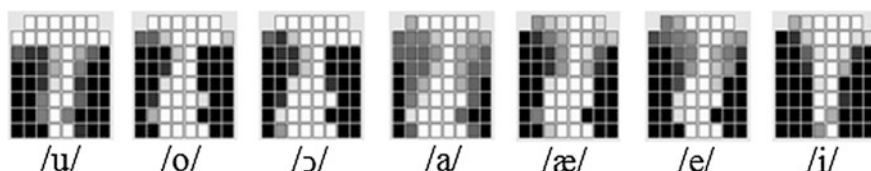
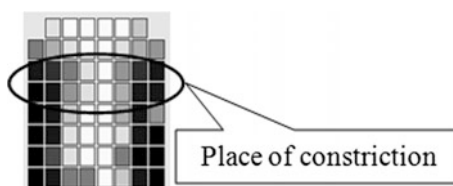


Fig. 2.83 Place of constriction of [tʃ] for seven Bangla vowels of male speaker

Fig. 2.84 Overall place of constriction of [tʃ]



Though there is no mentionable sex-wise difference in the place of articulation the male speaker show greater amount tongue-palate contact. The place of constriction is post alveolar.

Thus for this affricate occlusion occurs at alveolar position and constriction occurs at post alveolar position.

Average occlusion duration is 41.98 ms and average occlusion period energy is -42.51 dB. The length of the VOT (3.66 ms) is insignificant. Also there is no acoustic evidence of voicing. So it can be concluded that the Bangla consonant [tʃ] is unvoiced unaspirated affricate where the place of occlusion is alveolar and place of friction is post-alveolar.

Affricate ʈ [tʃʰ]

Figures 2.85 and 2.86 represents the overall place respectively of release and of the friction for the fricative [tʃʰ]. The place of release is seen as alveolar and that of friction is post-alveolar.

Figure 2.86 represents the overall place of friction of the fricative [tʃʰ], which is post-alveolar.

Average occlusion duration is 35.85 ms and average occlusion period energy is -42.94 dB. The length of the VOT is 27.23 ms large enough to be cognitively significant. There being no evidence of the glottal vibration it can be concluded that

Fig. 2.85 The overall place of release of [tʃʰ]

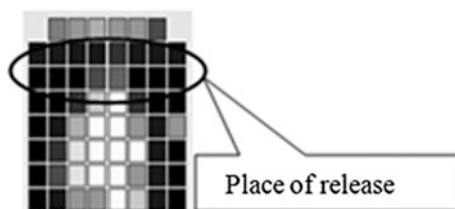
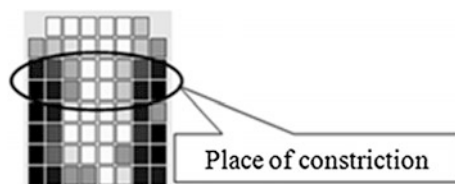


Fig. 2.86 The overall place of friction of [tʃʰ]



this Bangla consonant [tʃʰ] is unvoiced aspirated affricate where the place of occlusion is alveolar and place of friction is post-alveolar.

Affricate ত্র [dʒ]

Figures 2.87 and 2.88 represent the overall places respectively of release and constriction for the fricative [dʒ]. The place of release and friction is seen to be respectively at alveolar and post-alveolar region.

Average occlusion duration is 41.63 ms and average occlusion period energy is -30.07 dB. The average length of VOT (5.08 ms) is insignificant. This with the glottal vibration during the occlusion period indicates that the Bangla consonant [dʒ] is voiced unaspirated affricate where the place of articulation of occlusion is alveolar and place of articulation of friction is post-alveolar.

Fig. 2.87 Overall place of release [dʒ]

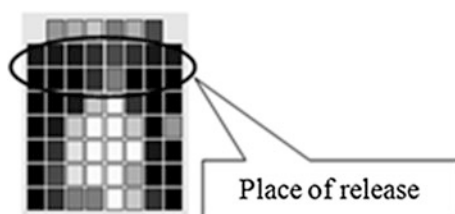


Fig. 2.88 Overall place of constriction [dʒ]

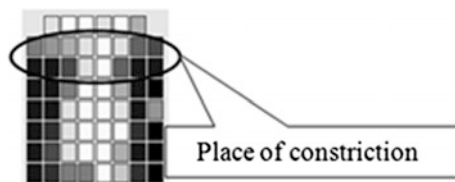


Fig. 2.89 Overall place of release of [ɖʱ]

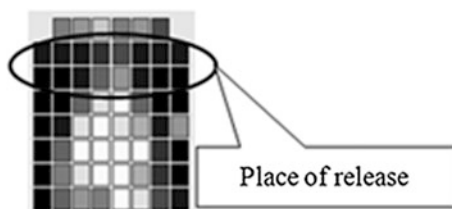
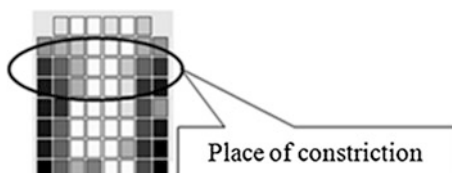


Fig. 2.90 Overall place of constriction of [ɖʱ]



Affricate ঙ [ɖʱ]

Figures 2.89 and 2.90 represents the overall place of release of [ɖʱ] and place of friction respectively

As the place of release and place of friction together represents the place of articulation then the place of articulation of release is alveolar and place of articulation of friction is post alveolar. Average occlusion duration is 34.04 ms and average occlusion period energy is -32.51 dB. From the acoustic evidence it can be stated that during the occlusion period the glottal vibration do exist. There is also an aspiration after the burst along with the voiced signal. The average length of the aspiration part is 26.02 ms. Thus the Bangla consonant [ɖʱ] is voiced aspirated affricate where the place of articulation of occlusion is alveolar and place of articulation of friction is post-alveolar. From detail vowel-wise EPGs for Bangla affricates it is noticed that the tongue tip is little bit forwarded during the release for front vowels.

Thus Bangla affricates, irrespective of their manner of production are alveolar as stop and post alveolar as release. This indicates that during the production of these sounds the tongue is retracted by one place of articulation.

2.5.4 Laterals

The place of articulation of the lateral is determined by the EPG frame collected from the middle of the production of the consonant.

Lateral ল [l]

Figures 2.91 and 2.92 represent the EPG frames collected at middle of the production of the consonant [l] with all the seven vowels spoken by the female and male speakers respectively. Let us peruse the EPG frame corresponding to the

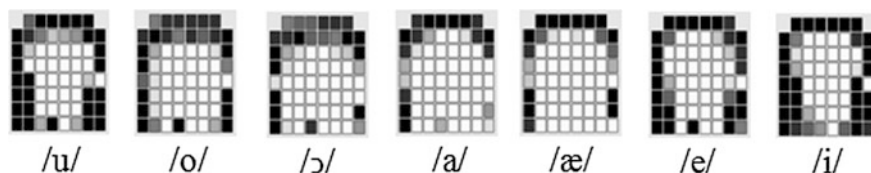


Fig. 2.91 EPG frame for all the seven vowel of female speaker

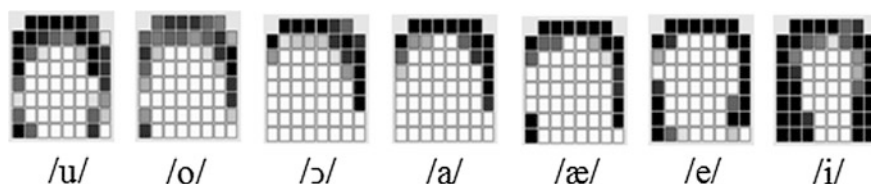
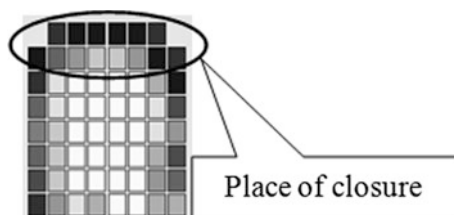


Fig. 2.92 EPG frame for all the seven vowels of male speaker

Fig. 2.93 Overall articulatory contact positions



vowel /a/, one may note here that it is the gray level representation of the ten repetitions. There is a clear lateral path on the right flank for all utterances as indicated by the three white cells on the right. At the same times there are two gray cells on the left flank which indicates that for some utterance this pronunciation is bilateral. Thus from minute examination of the average frames one can gather the nature and the side of lateralization for the laterals.

Figure 2.93 represents overall articulatory contact positions for the consonant [l].

From the above figures it is observed that during the production of consonant [l] the articulator completely closes the air passage medially at front and air passes thorough the side of the tongue. If the air passes through the one side of the tongue then it is called unilateral and if the air passes through the both side then it is called bi-lateral. From the above figures one could say that the Bangla consonant [l] is predominantly unilateral with the caution that only two informants have been examined. The place of articulation is dental. The average duration of the [l] in Bangla is 54.02 ms.

2.5.5 Nasal Murmur

The place of articulation of the murmur is determined by the EPG frame collected from the middle of the production of the consonant.

Consonant ঙ [ŋ]

Figure 2.94 represents overall articulatory closure positions for the consonant /ŋ/.

From the detailed vowel wise figure given in the Appendix it is seen that a complete closure is observed in case of front vowels. Due to the limitation of artificial palate sensor position as mention in Sect. 2.5.1 the complete closure is not observe in case of rest of the vowels. From Fig. 2.99 the place of closure is in the velar region. Therefore the place of articulation of [ŋ] is velar. The average duration of the [ŋ] in Bangla is 50.68 ms

Consonant ঞ [ɲ]

The occurrence of this consonant is rare. Only two Bangla words (মিঞা, যাজ্ঞা) has been used for the study of this consonant. The 5 repetitions of those words are recorded and the EPG frame collected from the middle of the production of the consonant is considered for the place of articulation determination.

Figure 2.95 represents the overall place of closure for the above consonant.

From the figure it is observed that the place of articulation of the above consonant is alveolar.

Consonant ণ [ɳ]

In general the [ɳ] is not pronounced in isolation in Bangla. It is only pronounced in cluster with retroflex oral sounds. Due to the above reason this consonant is recorded within a word where it is pronounced instead of usual VCV combination.

Fig. 2.94 Overall articulatory closure positions

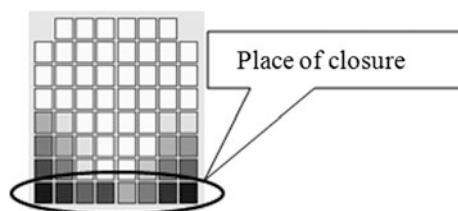


Fig. 2.95 Overall place of closure for /ɲ/

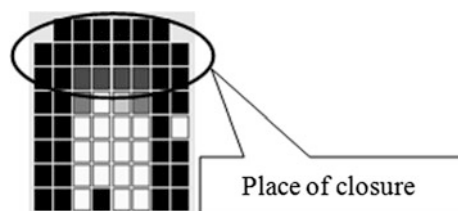


Fig. 2.96 Overall articulatory closure positions

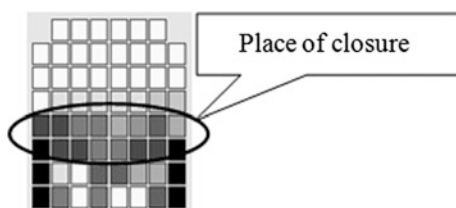
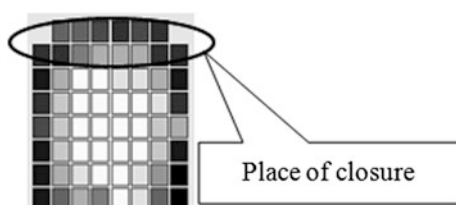


Fig. 2.97 Overall articulatory closure positions



EPG frames collected from the middle of the production of the consonant are considered for the determination of the place of articulation.

Figure 2.96 represents the overall EPG closure position for the above consonant.

From the above figure it is observed that the above consonant place of articulation is palatal. Its manner of production is retroflex nasal. The average contact area is 32.57%. The average duration of the [ŋ] in Bangla is 44.12 ms.

Consonant ण [n]

Figure 2.97 represents overall articulatory closure positions for the consonant [n].

The above figure shows that the place of contact this consonant is at the root of the teeth. Taking note of the deficiency of the palate in not having contacts on the teeth area, one may consider that the place of articulation of the above consonant as **dental**. The average duration of the [n] in Bangla is 54.65 ms.

Consonant ম [m]

After analyzing the EPG frame data it is observed that there is no closure in the whole palate region. The video of lips movement of the informant is recorded during the production of this sound and from the video evidence the closure is found in the lips region. Figure 2.98 represents the lips closure of during the production of the above sound for all the Bangla vowels. The figures show that the consonant [m] is bilabial. The average duration of the /m/ in Bangla is 57.06 ms

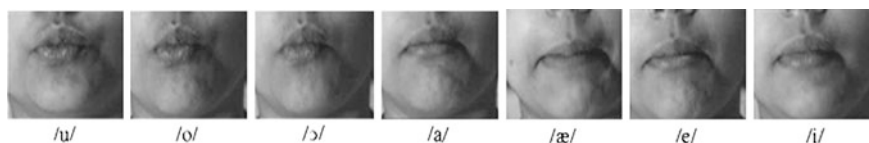


Fig. 2.98 Lips closure of during the production

In this experimental setup nasal spirometers to measure air passing out of nose was not available. Therefore the nasality of this group of consonants was determined from the acoustic evidence and subjective listening.

2.5.6 Trills, Flap or Tap

Consonant $\Re$ [r]

Figure 2.99 represents overall articulatory positions for the consonant [r].

The detail vowel wise frames show that there is no complete closure region. This may be due to the situation that [r] being found mostly a tap (see Sect. 2.5.6) the momentary contact of the tongue tip may occasionally miss capture in the EPG system. Therefore the continuous black line indicating complete closure is not always visible. However the gray level connectivity reflects the occasional capture of the blockage. Overall place of closure can be considered at alveolar region. So the places of articulation of the Bangla trill [r] is **alveolar**.

From an examination of the detailed vowel-wise EPG it is observed that the place of contact of the tip of the tongue is more forwarded in case of front vowel for both the speaker.

Consonant Ψ ([ɽ])

The place of articulation of these sounds is determined by the EPG frame collected from the middle of the production of the contact area. Contact area is measured by the number of contacts made to the total number of electrodes in the whole EPG frame and is expressed as percentage.

Figure 2.100 represents overall articulatory positions for the consonant [ɽ].

Fig. 2.99 Overall articulatory positions

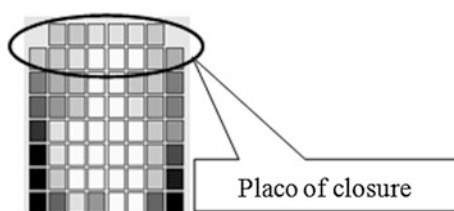


Fig. 2.100 Overall articulatory position of ([ɽ])

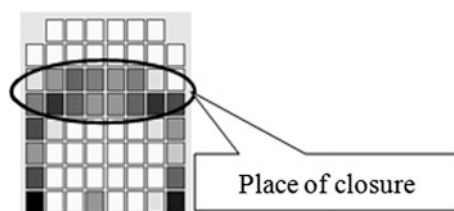
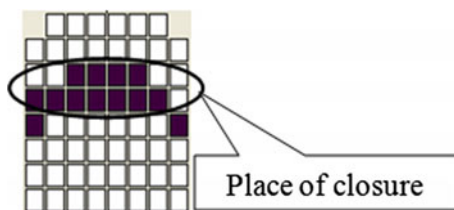


Fig. 2.101 Overall articulatory position of [tʰ]



The closure is in the region of post alveolar. The above consonant is retroflex flap and the place of articulation is **Post alveolar**.

Consonant ʈ ([ɽʰ])

The consonant [ɽʰ] is the aspirated counterpart of the consonant [ɽ], the place of articulation remaining the same as for [ɽ]. But the only difference is that after the release there is a perceptible amount of aspiration. Figure 2.101 represents the overall EPG data for the contact position for [ɽ]. Thus [ɽʰ] is post alveolar aspirated retroflex tap.

The EPG analysis confirms the traditionally known places of articulation for [ɽ, ɽʰ] and [ɽʰ]. However both EPG and acoustic analysis revealed the possibility of [ɽ] being a tap and [ɽʰ] being a trill, which is a deviation from the commonly held belief.

2.5.7 Area of Obstruction

The consonants dealt with above are some form of obstruent where sometimes firm and complete obstruction takes place and other times narrow constriction are made which produced quasi-random signals. The role of area where obstruction of the passage occurs needs some understanding regarding the role of articulators in the production. It has been noticed from detailed vowel-wise EPGs that the position as well as the contact area is influenced by the adjoining vowels. The two important measures are area of actual contact as well as area of constriction. Figure 2.102 presents the average contact areas of appropriate types for different consonants. Velar stops are not included because the artificial palates do not cover the appropriate region. Similarly bilabials are not included because of obvious reasons. Taps and trills do not exhibit firm contacts.

A perusal of the figures reveals that the area of contact/constriction depends consistently on the adjoining vowel. The area is least for central low vowel [a]. The area rises from low to high. For low vowels the vertical movement is larger and therefore one may expect lesser time for the contact with the palate. This could be a possible explanation. Further there is a general trend of increase in the area from back to front, very sharply towards front. This is expected as with the front vowels the contact for the consonant is made by raising the tip of the tongue while for the vowel the portion near the root is raised. This indicates a complex motion for the

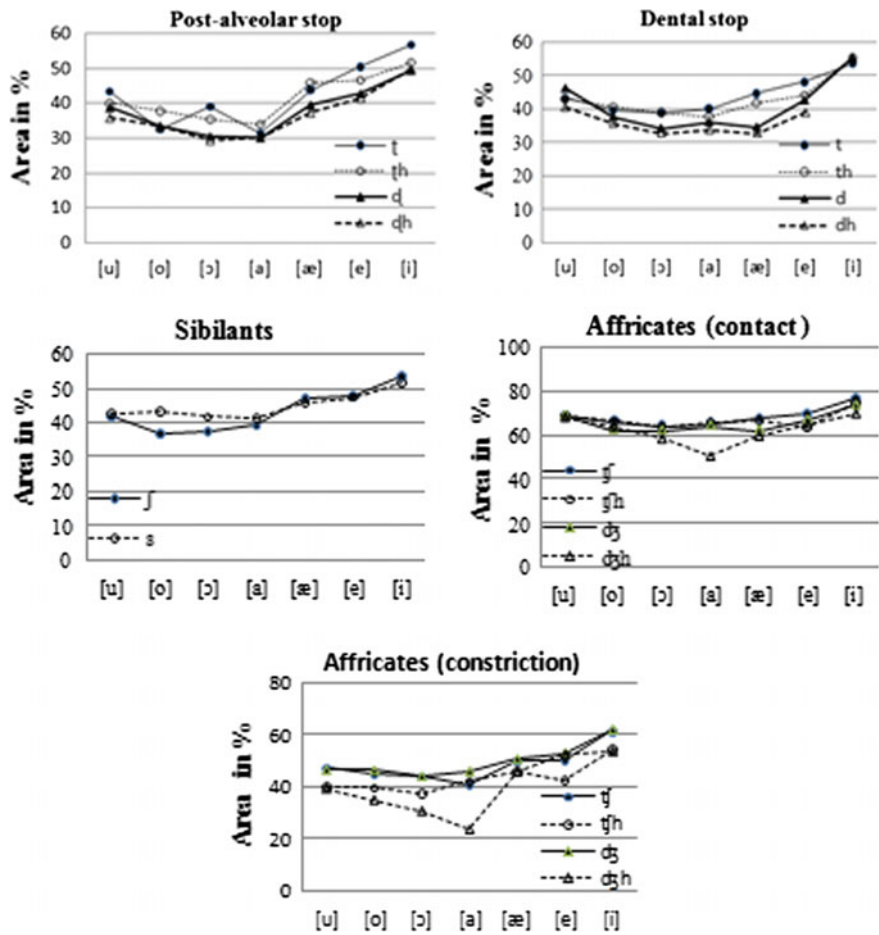


Fig. 2.102 The area of contact/constriction in the oral tract for different consonants

tongue for front leaving lesser time for the consonantal contact. As far as the articulatory positions of stops are concerned both the magnitude and the range of contact area increases from back to front. Area of constriction for affricates is consistently lower than that for firm closure.

2.5.8 Summary

For plosive/stops there is a general indication that contact region gets generally forwarded on release by about one row, which is not very insignificant. It indicates that the closure is released by a forward movement of the tongue. For retroflex

place of release advances two rows taking it to Post alveolar region from a mid-palatal closure. The dependence of the place of articulation with respect to sex is consistent but quite moderate, generally forwarded for the female speaker. The influence of vowels on place of release and closure is significant. It is monotonically forwarded from back to front, generally two rows, most backward for /u/ and most forward for /i/.

It may be noticed from Fig. 2.1 that in EPG frame configuration a single row difference may sometime indicate different category in articulation. Therefore identifying plosives/stops by the present system of fine categorization may sometimes be misleading. For other sounds the problem is not significant.

The EPG provides for the first time a vivid visual presentation of the dynamic process of the tongue-body interaction in the articulation of consonants. Though it is intermittent it provides an unambiguous picture within its stated limitation. This provides a new understanding that the place of articulation is not so crisp a categorization as one would like to believe from the traditional description and IPA symbolic representation. In fact the fuzziness is so large in the case of retroflex plosives/stops that one has to reconsider how to define them. The release is at one place and the closure is at a different categorical place. Also the extent of influence of vowels on deciding the place category is too significant to be wished away. A wholesome picture would come out only when this type of study is done with larger number of informants.

One of the important shortcomings of the EPG system is that it does not provide firm contact evidence for velar and dental region. The other one is that as of now it is difficult to have studies done with a large number of speakers. The false palate has to be made outside India increasing the cost and time. Once this is made in India larger number of informants could be used. I am sure necessary technological is available in the country. Only one has to tap them.

References

- Andruski JE, Ratliff M (2000) Phonation types in production of phonological tone the case of Green Mong. *J Int Phon Assoc* 30:39–62
- Berkins S, Stevens KN (1982) Across language study of the perception of nasal vowels. *J Acoust Am* 73(suppl 1):s54
- Bhattacharya T (2000) Bangla (Bengali). In: Gary J, Rubino C (eds) *Encyclopaedia of world's languages: past and present*. WW Wilson, New York
- Bickley C (1982) Acoustic analysis and perception of breathy vowels. In: *Speech communication group working papers*, pp 71–82. Cambridge Massachusetts Institute of Technology
- Catford JC (1977) *Fundamental problems in phonetics*. Indiana University Press
- Chatterji SK (1926) *The origin and development of the Bengali language*. Rupa & Co., New Delhi
- Chávez-Peón ME (2013) *Non-modal phonation in Quiavini Zapotec: an acoustic investigation*. Instituto de Investigaciones Antropológicas, Universidad Nacional Autónoma de México. Retrieved 26 May 2013
- Choudhury S (2006) *Concatenative text to speech synthesis: a study of Standard Colloquial Bengali*. PhD Thesis, Indian Statistical Institute

- Colin PM (1991) Cited in Indo-Aryan languages. Cambridge University Press
- Crystal D (2003) A dictionary of linguistics & phonetics (5th ed). Wiley-Blackwell
- Das Mandal SK (2007) Role of shape domain parameters in speech recognition: a study on Standard Colloquial Bangla (SCB), Ph.D Thesis, Jadavpur University
- Datta AK (1988) Acoustic phonetics of non-nasal standard Bengali vowels: a spectrographic study. *JIETE* 34
- Datta AK (2014) Aspirated vowels in Standard Colloquial Bengali: a case study with native informants' communicated to JPH, Springer, Berlin
- Datta AK, Ganguly NR (1981) Terminal frequencies in CV Combination in multisyllabic words. *Acustica* 47(4):314–324
- Datta AK, Ganguly NR (1985) Behaviour of terminal frequencies in VC combination. *Acustica* 67:26–33
- Datta AK, Mukherjee B (2011) On the role of formants in the cognition of vowels and place of articulation of plosives. In: Ystad S, Aramaki M, Kronland-Marinet R, Mohanty S (eds) *Speech, sound and music processing: embracing research in India*
- Datta AK, Dutta Majumder D, Ganguly NR, Mukherjee B, Sarkar R (1974) Studies on acoustic phonetic features of Telugu speech sounds. ECSL series on phonological studies, Nov 1974
- Datta AK, Ganguly NR, Ray S (1978a) Transition—a cue for identification of plosives. *J Acoust India* VI(4):124–131
- Datta AK, Ganguly NR, Dutta Majumder D (1978b) Some studies on acoustic features of Telugu vowels. *Acustica* 41:55–64
- Datta AK, Ganguly NR, Ray S (1980) Recognition of unaspirated plosives: a statistical approach. *IEEE Trans Acous Speech Sig Process ASSP-28*(1):85–91
- Datta AK, Ganguly NR, Dutta Majumder D (1981) Acoustic features of consonants; a study based on Telugu speech sounds. *Acustica* 47(2)
- Datta AK, Ganguly NR, Mukherjee B (1989) Bengali nasal sounds—a spectrographic study. *J Acous Soc India* XVII:219–223
- Datta AK, Sengupta R, Dey N, Banerjee BM, Nag D (1998) Perception of nasality in Bengali vowels: role of harmonics between F0 and F1. In: *Proceedings of international conference of computational linguistics, speech and document processing, ISI, Calcutta, Feb 18–20, 1998*
- Datta AK, Sengupta R, Dey N (2003) Jitter, Shimmer and HNR characteristics of singers and non-singers. *J ITC Sangeet Res Acad* 17
- Delattre PC, Liberman AM, Cooper FS, Gerstman LJ (1952) An experimental study of the acoustic determinants of vowel colour; observations on one- and two-formant vowels synthesized from spectrographic patterns. *Word* 8(3):195–210
- Dioubina OI, Pfitzinger HR (2002) An IPA vowel diagram approach to analysing L1 effects on vowel production and perception. In *Proceedings of ICSLP '02, Denver, vol 4, pp 2265–2268*
- Djordje K, Das R (1972) Short outline of Bengali Phonetics. Statistical Publishing Society, Calcutta
- Dutta Majumder D, Datta AK (1966) A scheme for automatic speech coding and recognition. In: *ISALS symposium on control computation*
- Esposito CM, Khan SUD, Hurst A (2005) A breathy nasals and /Nh/Clusters in Bengali, Hindi, and Marathi. *UCLA working papers in phonetics, vol 104, pp 82–106*
- Fant G (1970) Acoustic theory of speech production. Mouton De Gruyter
- Fant G, Ishizaka K, Lindquist-Gauffin J, Sundberg J (1972) Subglottal formants, STL-QPSR 1/1972
- Fischer-Jorgensen E (1967) Phonetic analysis of breathy (murmured) vowels in Gujarati. *Indian Linguist* 28:71–139
- Fujimara O (1960) Spectra of nasalised vowels. *Res Lab Electr Q Prog Rep No. 58, MIT* 214–218
- Ganguly NR, Datta AK, Mukherjee B (1988) Acoustic phonetics of non-nasal Standard Bengali Vowels: a spectrographic study. *J Electron Telecommun Eng* 34(1):50–56
- Ganguly NR, Datta AK, Mukherjee B (1999) Acoustic phonetic features of glides and diphthongs in Bengali. *J Acous Soc India* XXVII:199–202

- Gordon M (2001) Phonation types: a cross-linguistic overview. *J Phon*
- Hai AM (1989) Dhani Bigyan and Bangla Dhani Tattwa. Mallick Brothers, Dhaka, Bangladesh
- Hanson H (1995) Glottal characteristics of female speakers, Ph.D. dissertation. Harvard University, MA
- Hartmut RP (2003) Acoustic correlates of the IPA vowel diagram. 15 ICPHs Barcelona, pp 1441–1444
- Hawkins S, Stenens KN (1985) Acoustic and perceptual correlates of the non-nasal-nasal distinction for vowels. *J Acoust Am* 77(4):1560
- Heinz JM, Stevens KN (1961) On the properties of voiceless fricative consonants. *J Acoust Am* 33:589–596
- Hideaki K, Hideo K, Jun T, Masaru S (2006) Spectral properties of Japanese whispered vowels referred to pitch. *J Acoust Soc Am* 120(5):3378–3378
- Hillenbrand J, Houde RA (1996) Acoustic correlates of breathy vocal: quality dysphonic voices and continuous speech. *J Speech Hear Res* 39:311–321
- Hillenbrand J, Cleveland RA, Erickson RL (1994) Acoustic correlates of breathy vocal quality. *J Speech Hear Res* 37:769–778
- Hombert JM, Ohala J, Ewan W (1979) Phonetic explanations for the development of tones. *Language* 55(1):37–58
- Hossain SA, Rahman ML, Ahmed F (2005) Acoustic space of Bangla vowels. In: WSEAS 5th international conference on speech and image processing, Greece, pp 138–140, August 2005
- Huffman MK (1987) Measures of phonation types in Hmong. *J Acoust Soc Am* 81(1):495–504
- Ioana C (2002) A perception-production study of Romanian diphthongs and glide-vowel sequences. *J Int Phon Assoc* 32(2):203–221
- Jones D (1962) An outline of English phonetics. W. Heffer & Sons Ltd., Cambridge
- Klatt DH, Klatt C (1990) Analysis, synthesis, and perception of voice quality variations among female and male talkers. *J Acoust Soc Am* 87(2):820–857
- Ladefoged P (1967) Three areas of experimental phonetics. Oxford University Press, Oxford
- Ladefoged P, Antananzas-Barroso N (1985) Computer measures of breathy voice quality. UCLA working papers in phonetics, vol 61, pp 79–86
- Ladefoged P, Maddieson I (1996) The sounds of the world's languages. Blackwell, Oxford
- Cambridge, MA
- Laver J (1980) The phonetic description of voice quality. Cambridge University Press
- Lehiste I, Peterson EG (1961) Transitions, glides, and diphthongs. *JASA* 33(3):268–277
- McGowan RS (1922) Tongue-tip trills and vocal-tract wall compliance. *JASA* 91:2903–2910
- Mistry PJ (1997) Gujarati phonology. In: Kaye AS (ed) Phonologies of Asia and Africa. Winona Lake Eisenbrauns
- Ngoc YPT, Badin P (1994) Vocal tract acoustic transfer function measurements further developments and applications. *Suppliment au Journal de Physique* 111, vol 4, May 1994
- Padgett J (2007) Glides, vowels, and features. *Lingua* 118(12):1937–1955
- Pandit PB (1954) Indo Aryan sibilants in Gujarati. *IL*, 14
- Perkell JS (1969) Physiology of speech production: results and implications of a quantitative cineradiographic study. Research Monograph No. 53, MIT Press, Cambridge MA
- Recasens P, Pallare MD (1999) A study of /Q/ and /r/ in the light of the “DAC” coarticulation model. *J Phon* 27:143–169
- Saksena BR (1971) *Evolution of Awadi*. Motilal Banarasi Das Publication, New Delhi, pp 74–76
- Sarkar P (1985) Bangla Dwishardhani (Bangla Diphthong). Calcutta, 1985–86
- Schane S (1995) Diphthongization in particle phonology. In: Goldsmith JA (ed) The handbook of phonological theory. Blackwell handbooks in linguistics. Blackwell, pp 586–608
- Shaddle CH (1990) Articulatory-acoustic relationships in fricative consonants. In: Hardcastle WJ, Marchal A (eds) Speech production and speech modelling. Kulwer, Dordrecht, The Netherlands, pp 187–209
- Spajić S, Ladefoged P, Bhaskararao P (1996) The trills of Toda. *J Int Phon Assoc* 26(1):1–22
- Stevens K (2000) Acoustic phonetics. MIT Press, Massachusetts

- Stevens K, Hanson H (1994) Classification of glottal vibration from acoustic measurements. Presented at the 8th vocal fold physiology conference, Kurume, Japan, April 7–9, 1994
- Takeuchi S, Kasuya H, Kido K (1975) On the acoustic correlates of nasality. *J Acoust Jpn* 31: 298–309
- Tartter VC (1989) What's in a whisper? *J Acoust Soc Am* 86(5):1678–1683
- Teagre HM, Teagre SM (1990) A phenomenological model for vowel production in the vocal tract. In: Daniloff RG (ed) *Speech science recent advances*, College Hill, San Diego, pp 73–109
- Thongkum T (1988) Phonation types in Mon-Khmer languages. In: Fujimura O(ed) *Vocal fold physiology voice production, mechanisms and functions*. Raven Press, New York, pp. 319–334
- Tsunoda K, Ohta Y, Soda Y, Niimi S, Hirose H (1997) Laryngeal adjustment in whispering: magnetic resonance imaging study. *Ann Otol Rhinol Laryngol* 106:41–43
- Varshney RL (1995) *An introductory textbook of linguistics and phonetics*, 8th edn. Student Publication store
- Wayland R, Jongman A (2003) Acoustic correlates of breathy and clear vowels the case of Khmer. *J Phon* 31:181–201
- http://www.cdackolkata.in/html/txttospeech/corpora/corpora_main/MainB.html
- <http://www.phonetics.ucla.edu/course/chapter1/vowels.html>
- <http://www.speech.kth.se/wavesurfer>

Acoustics of Bangla Speech Sounds

Datta, A.K.

2018, XIV, 133 p. 210 illus., 137 illus. in color.,

Hardcover

ISBN: 978-981-10-4261-4